

TECHNISCHE UNIVERSITÄT ILMENAU

Fakultät für Elektrotechnik und Informationstechnik

Bachelorarbeit

Erstellung einer Datenbank von kopfbezogenen Impulsantworten

vorgelegt von:	Mathias Hellmich
geboren am:	31. 10. 1987 in Dresden
Studiengang:	Medientechnologie

Verantwortlicher Hochschullehrer:	Prof. Dr.-Ing. Dr. rer. nat. h.c. Brandenburg
Wissenschaftlicher Betreuer:	Dipl.-Ing. Stephan Werner

Beginn der Arbeit:	15.09.2011
--------------------	------------

Arbeit eingereicht am:	09.03.2012
------------------------	------------

Ilmenau, 09.03.2012

Danksagung

Ich bedanke mich ganz herzlich bei allen Personen, die diese Arbeit ermöglicht haben. Bei meinem Betreuer Stephan Werner und seinem Kollegen Florian Klein für ihre Verbesserungsvorschläge und die konstruktive Kritik. Ich bedanke mich bei meinen Eltern und ihrem unermüdlichen Einsatz, meine Rechtschreibfehler zu korrigieren. Ebenso möchte ich Ming Liu für ihre Unterstützung danken.

Kurzfassung

Dreidimensionale Wiedergabe von Bild und Ton sind derzeit einige der wichtigsten Forschungsfelder für die Medienwiedergabe. Um eine binaurale Tonwiedergabe, mit der ein dreidimensionales Tonerlebnis erzeugt wird, ermöglichen zu können, werden kopfbezogene Übertragungsfunktionen eines Kunstkopfes aufgenommen und in einer Datenbank gespeichert. Diese Funktionen beschreiben sehr präzise die Richtungsmerkmale von Schallquellen anhand von nur zwei Wiedergabequellen. Die Wiedergabe findet dabei über Kopfhörer statt und ist stark von den individuellen Hörern abhängig.

Diese Arbeit beinhaltet die Grundlagen, auf denen die binaurale Wiedergabe basiert. Sie befasst sich mit binauralen Merkmalen sowohl im allgemeinen als auch speziell bei kopfbezogenen Übertragungsfunktionen. Ebenso werden verschiedene Mikrofonierungsverfahren angesprochen. Es wird eine Datenbank von kopfbezogenen Übertragungsfunktionen anhand eines Kunstkopfes aufgenommen. Die Durchführung des Vorhabens wurde mitsamt dem Aufbau und den auftretenden Problemen erläutert.

Das Endergebnis dieser Arbeit ist eine umfangreiche und möglichst genaue Datenbank aus kopfbezogenen Übertragungsfunktionen für die horizontale Ebene. Ein Vergleich zu einer ausgewählten Messung mit einem Probanden wurde durchgeführt und analysiert, um eine Grundlage für die Erstellung einer individualisierten Datenbank zu liefern. Die entstandene Datenbank soll später sinnvoll zu weiteren Forschungszwecken eingesetzt werden können. Die akustischen Merkmale wurden getrennt gespeichert, um eine Laufzeitanpassung an individuelle Frequenzkomponenten durchführen zu können.

Abstract

Three-dimensional reproduction of images and sounds are currently some of the most important research fields for media development. To create a binaural sound, with which a three-dimensional sound experience is made possible, head-related transfer functions are taken from a dummy head and stored in a database. They describe very accurately the direction characteristics of sound sources using only two playback sources. These two source are played via headphones and are highly dependent on individual listeners.

This paper deals with the basics, which make the binaural reproduction possible. It furthermore deals with binaural features both generally and specifically for the head-related transfer functions. Also addressed are various miking processes. The implementation of the project was discussed along with the construction and the problems that arise.

The result of this work is a comprehensive and most accurate database of head-related transfer functions in the horizontal plane. Comparison between the used Dummy Head and one proband were executed and analysed in order to provide a basis for the creation of an individualized database. The resulting database is determined to be used for research purpose later on. The acoustic features were stored separately in order to perform a run-time adaptation to individual frequency components.

Inhaltsverzeichnis

Kurzfassung.....	iii
Abstract	iv
1. Einleitung.....	1
2. Grundlagen.....	2
2.1 Akustische Merkmale.....	2
2.1.1 Interaurale Laufzeitdifferenz.....	3
2.1.2 Interaurale Pegeldifferenz.....	4
2.1.3 Interaurale Korrelation.....	6
2.1.4 Blauertsche Bänder.....	7
2.2 Head-Related Transfer Function.....	8
2.2.1 Definition und Erklärung.....	9
2.2.2 Eigenschaften.....	10
2.2.3 Binaurale Merkmale bei HRTF's.....	11
2.2.4 Individualisierte HRTF.....	12
2.2.5 Verwendung/Funktionsweise HRTF.....	13
2.3 Stereophonieverfahren.....	15
2.3.1 Differenzstereophonie.....	15
2.3.2 Äquivalenzstereophonie.....	16
2.3.3 Trennkörperstereophonie.....	17
2.3.4 Kugelflächenstereophonie.....	17
2.3.5 Kunstkopfstereophonie.....	18
2.3.6 Head and Torso Simulator.....	19
2.4 Raumimpulsmessungen.....	20
2.4.1 Klassische Messung.....	20
2.4.2 Sweep Verfahren.....	20
2.4.3 TDS-Verfahren.....	21
2.4.4 MLS-Technik.....	22
3. Durchführung.....	23
3.1 Verwendete Geräte.....	23
3.2 Messaufbau.....	24
4. Auswertung.....	26
4.1 Aufbereitung der Messergebnisse.....	26
4.1.1 Fensterfunktion.....	26
4.1.2 Differenzextraktion.....	28
4.1.3 Minimalphasen Approximation.....	29
4.2 Auswertung der Messergebnisse.....	31
4.2.1 Minimalphasenreduktion.....	31
4.2.2 Amplitudenspektrum.....	33
4.2.3 Probandenmessungen.....	34
5. Fazit und Ausblick.....	39
6. Abbildungsverzeichnis.....	41
7. Literaturverzeichnis.....	43
8. Abkürzungs - und Symbolverzeichnis.....	47
Ehrenwörtliche Erklärung.....	48

1. Einleitung

Ein weites Forschungsfeld in der Akustik ist die dreidimensionale Wiedergabe von Schall. Diese Wiedergabe kann bereits sehr gut über Kopfhörer ermöglicht werden. Dazu werden am Kopf des Menschen, der in dieser Arbeit durch einen Kunstkopf repräsentiert wird, Schallmessungen durchgeführt. Das Ergebnis ist eine Datenbank richtungsabhängiger Übertragungsfunktionen. Sie stellen das Verhalten von akustischen Signalen von der Signalquelle bis hin zum Ohreingang des Menschen dar und sind eine allgemeine Beschreibung für Merkmale zur Lokalisation, die international als Head-Related Transfer Function (HRTF) bekannt sind und mit kopfbezogene Übertragungsfunktionen übersetzt werden können. [Blauert, 1997]

In HRTF - Datenbanken werden die Richtungsinformationen verschiedener winkelbezogener Positionen in Bezug zu einem bestimmten Kopf aufgelistet und zur dreidimensionalen Darstellung verwendet.

Um eine Anpassung an einzelne Personen zu gewährleisten, wird im Umfang dieser Arbeit ein Kunstkopf hinsichtlich seiner HRTF vermessen. Die entstandenen Datenbanken werden so aufbereitet, dass man Pegel - und Laufzeitdifferenzen separat entnehmen kann und gegebenenfalls für eine Individualisierung verwenden kann. Damit wird eine Grundlage für weitere Forschungen anhand des Kunstkopfes geschaffen.

Der Umfang der Arbeit ist im folgenden aufgelistet.

- Messung der HRIR's an einem Kunstkopf im reflexionsarmen Raum,
- Auswertung und Visualisierung der gemessenen HRIR's,
- Auftrennung der ermittelten HRIR's in interaurale Laufzeit - und interaurale Intensitätskomponenten,
- Überprüfung und Vergleich der aufgetrennten HRIR Datensätze mit den ursprünglich gemessenen Datensätzen,
- Erstellung der HRTF Datenbank für die horizontale Ebene,
- Exemplarischer Vergleich der aufgenommenen Datenbank mit einer Probandenmessung,
- Beispielhafte Auswertung der Kunstkopfmessung und einer Probandenmessung.

2. Grundlagen

2.1 Akustische Merkmale

Akustische Merkmale sind die Grundbausteine der Akustischen Lokalisation. Sie verweisen auf akustische räumliche Merkmale, die für die Lokalisierung von Schall wichtig sind. Durch sie ist es dem Menschen möglich, den Ort eines akustischen Ereignisses relativ genau zu bestimmen. Auch in den Head - related Transfer Functions, die später noch genauer beschrieben werden, sind sie wichtiger Bestandteil.

Es wird bei den akustischen Merkmalen unterschieden zwischen monauralen (auf ein Ohr bezogen) und binauralen (auf beide Ohren bezogen) Merkmalen. Zu den binauralen Merkmalen gehören die interaurale Pegeldifferenz, die interaurale Laufzeitdifferenz, die interaurale Kreuzkorrelation und die spektrale Verteilung. Hingegen kann man als monaurales Merkmal nur das Frequenzspektrum des jeweiligen Ohres betrachten. Im nachstehenden Kapitel werden diese Merkmale näher erläutert.

Um eine dreidimensionale Beschreibung von Signalen zu erstellen, werden zwei Winkel eingeführt, wie in Abbildung 1 dargestellt, um sämtliche Richtungen definieren zu können: Der Azimuthwinkel φ , welcher die horizontale Auslenkung definiert und der Elevationswinkel ϑ , der zur Beschreibung der vertikalen Auslenkung dient. Zusammen mit diesen Winkeln werden drei Lokalisationsebenen definiert.

Die Horizontalebene, welche sich für den Betrachter horizontal auf Ohrhöhe befindet. Die Frontalebene ist eine Ebene, welche sich senkrecht zur Bezugsperson befindet und wieder durch beide Ohren verläuft. Die letzte Ebene nennt sich Medianebene und verläuft 90° versetzt zu den anderen beiden. Sie schneidet die beiden Ohren nicht, sondern verläuft durch Nase und Hinterkopf der Bezugsperson.

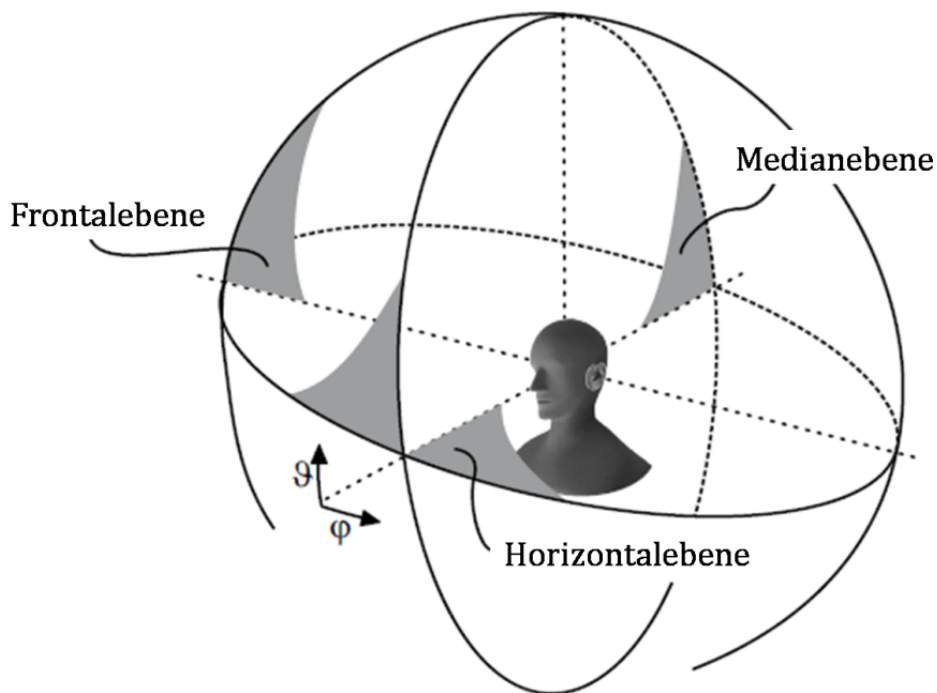


Abbildung 1: Kopfbezogenes, dreidimensionales Koordinatensystem mit Azimutwinkel φ und Elevationswinkel ϑ [Fels, 2008]

2.1.1 Interaurale Laufzeitdifferenz

Die interaurale Laufzeitdifferenz (engl.: Interaural Time Difference, ITD) ist eine physikalische Eigenschaft des Schalls, welche zur Lokalisation von akustischen Signalen in der Horizontalebene beiträgt.

Wenn ein Signal aus einer bestimmten Richtung mit einem Winkel φ ungleich Null zu einem Hörer gelangt, wird es zwangsläufig nicht bei beiden Ohren gleichzeitig wahrgenommen, sondern zeitlich ein wenig versetzt. Dieser zeitliche Versatz wird interaurale Laufzeitdifferenz genannt. Der Grund dafür ist, dass sich Schall in Luft mit einer gleichmäßigen Geschwindigkeit c ausbreitet ($c = 343 \text{ m/s}$ bei 20°C) und somit zu dem zweiten nur wenige Zentimeter weiter entfernten Ohr auch mehr Zeit benötigt. Die zusätzliche Schallstrecke variiert je nach Winkel φ und hat sein Maximum bei 90° bzw. 270° .

Der Durchmesser d des KEMAR DummyHead wird als der Durchschnitt des menschlichen Kopfes angenommen und beträgt somit 17 Zentimeter

[Kemar, 2008]. Aus $s = \frac{d \cdot \pi}{2}$ lässt sich der zusätzliche maximale Weg s des

Schalls berechnen, der genau der halbe Umfang des menschlichen Kopfes beträgt und einen Wert von etwa 27 Zentimetern annimmt. Aus der Geschwindigkeit c und der errechneten maximalen Zusatzstrecke s lassen sich

die Größenordnungen für die Laufzeitdifferenz t nach $t = \frac{s}{c}$ herleiten. Diese liegen zwischen 0 μ s und 800 μ s. Dies kann anhand einer Prinzipskizze in Abbildung 2 nachvollzogen werden. In dieser stellt d den Kopfdurchmesser dar, α den Winkelversatz des eintreffenden Schalls und Δs die Zusatzstrecke, woraus sich die ITD berechnen lässt.

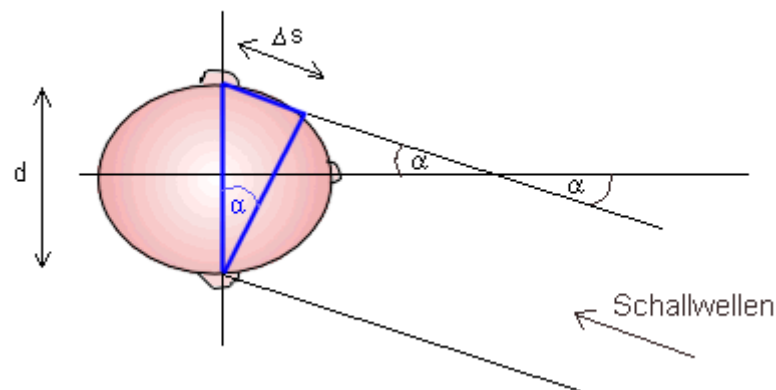


Abbildung 2: Prinzipskizze zur Bestimmung der Laufzeitdifferenz
[Kremer, 2012]

Da bei den Messungen nur eine Schallquelle verwendet wurde, lässt sich die interaurale Laufzeitdifferenz anhand der Differenz der Verzögerung der beiden Impulsantworten an den Ohren bestimmen, wie in späteren Kapiteln noch beschrieben wird.

Zusätzlich zur interauralen Laufzeitdifferenz wird ebenso die interaurale Phasendifferenz (IPD) genannt, da diese anhand der ITD ermittelt wird. [Yost, 1994]

2.1.2 Interaurale Pegeldifferenz

Wenn Schall von einer Quelle ausgesendet wird und bei den Ohren ankommt, wird neben der ITD ebenfalls die interaurale Pegeldifferenz, auch bekannt als Interaural Level Difference (ILD) wirksam. Diese zeigt auf, dass das

schallabgewandte Ohr einen wesentlich geringeren Schallpegel empfängt, als das dem Schall zugewandte Ohr.

Diese Pegelunterschiede sind ein Resultat von Reflexionen, Beugungen und Resonanzeffekten, welche auf den Kopf, den Oberkörper und auch der Form der Außenohren des Hörers zurückzuführen sind [Faller, 2004]. Zusätzlich dazu nimmt der Schalldruck und somit der Schallpegel für jede zurückgelegte Strecke ab. Diese Abnahme trägt allerdings nur sehr wenig zur interauralen

Pegeldifferenz bei. Durch das Entfernungsgesetz $\frac{p_2}{p_1} = \frac{r_1}{r_2}$, wobei p der Schalldruck ist und r der Abstand zur Schallquelle, lässt sich ein Pegelunterschied von maximal 2 dB errechnen.

Aus der Duplex Theorie nach Rayleigh geht die Erkenntnis hervor, dass ILD und ITD nicht einzeln der Lokalisation dienen, sondern eine Wichtung aus beiden Differenzen für den Menschen wichtig ist [Middlebrooks, 2002]. Dabei dominiert die Laufzeitdifferenz bei der Lokalisation niederfrequenter Töne, hingegen kommt die Pegeldifferenz bei hohen Tönen zum tragen.

Eine Skizze, welche den Übergang der Relevanz der interauralen Merkmale vermitteln soll, stellt Abbildung 3 dar.

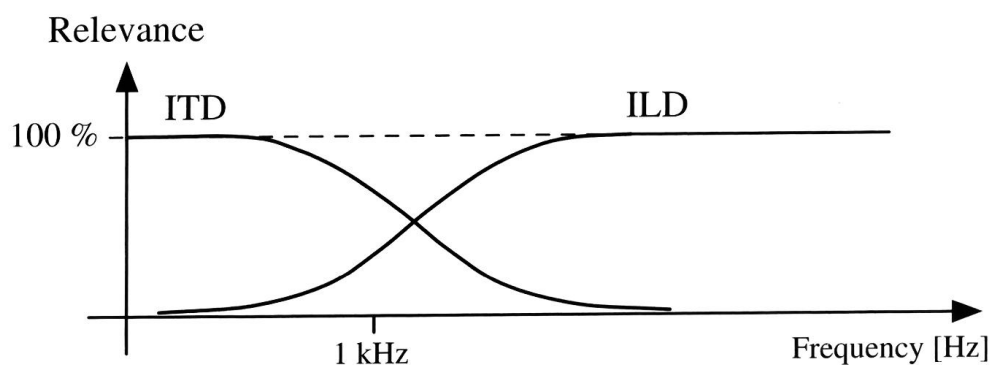


Abbildung 3: Relevanz der Differenzen anhand der Duplex Theorie [Breebart, 2007]

Dies wird erklärbar dadurch, dass niederfrequente Töne eine große Wellenlänge besitzen und ein Gegenstand von der Größe eines Kopfes kaum ein Hindernis für die Ausbreitung dieser Wellenlänge darstellt, sondern um das Hindernis gebeugt werden kann [Fedke, 2007]. Pegeldifferenzen würden also kaum zum Tragen kommen, Zeitdifferenzen würden allerdings ohne Probleme

erkennbar sein. Bei hohen Tönen mit kleiner Wellenlänge können jedoch Pegelunterschiede von bis zu 35 dB auftreten, die die interaurale Pegeldifferenz für die Lokalisation prädestinieren. [Breebart, 2007]

2.1.3 Interaurale Korrelation

Die interaurale Korrelation (IC), eine Kurzform des interauralen Kreuzkorrelationskoeffizienten (IACC), dient als drittes Modell des menschlichen Hörens der Erklärung und Beschreibung der Räumlichkeit eines Signals. Sie ist. [Möser, 2009]

Der IACC ist ein binaurales, kopfbezogenes Kriterium, welches der Beschreibung der Ähnlichkeit zweier Signale dient. Die zwei Signale hierbei sind die Raumimpulsantworten der beiden Ohrsignale, die frequenzgefiltert korreliert sein können. Der IC ist dabei das Maximum der normierten Kreuzkorrelationsfunktion in frei wählbaren Zeitgrenzen t_1 und t_2 :

$$IACC = \max \left| \frac{\int_{t_2}^{t_1} p_L(t) \cdot p_R(t+\tau) dt}{\sqrt{\left(\int_{t_2}^{t_1} p_L^2(t) dt \cdot \int_{t_2}^{t_1} p_R^2(t) dt \right)}} \right| \text{ für } -1 \text{ ms} < \tau < 1 \text{ ms} \quad [1]$$

Dabei stellen p_L und p_R jeweils den linke und rechte Ohrsignal dar. Die Länge der zeitlichen Verschiebung, über welchen die Korrelation durchgeführt wird, wird mit τ angegeben.

Mit dem IC wird allerdings vielmehr die „Breite“ eines Signals erklärt, also ob ein Signal einer Punktquelle oder eher einer räumlich ausgedehnten Quelle entspringt. Bei einer Punktschallquelle unterscheiden sich die Signale an beiden Ohren kaum, was in einem Korrelationskoeffizienten von „1“ resultiert. Dieser wird kleiner, je großflächiger das Quellensignal wird und resultiert in 0, wenn zwei separate Signale vorhanden sind. Dies wird in Abbildung 4 veranschaulicht, welche darstellt, wie man eine Signalbreite mithilfe von zwei Lautsprechern und dem IC simuliert. [Breebart, 2007; Möser, 2009]

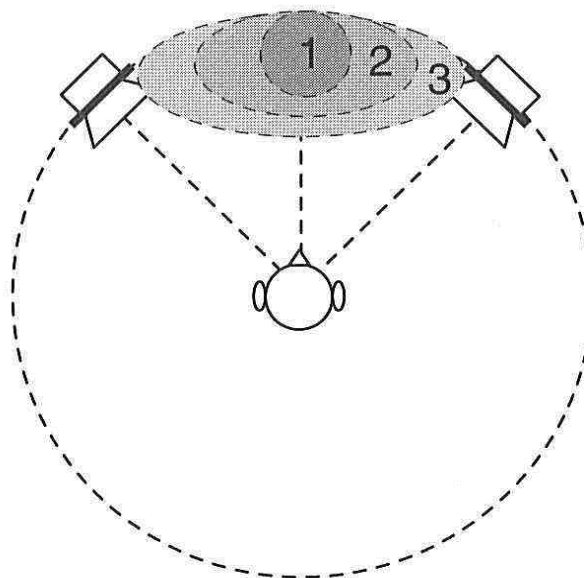


Abbildung 4: Demonstration der Breite eines akustischen Objektes anhand einer Stereolautsprecheranordnung, Breite erhöht sich (1-3), während sich die IC zwischen den beiden Lautsprechern verringert. [Breebart, 2007]

2.1.4 Blauertsche Bänder

Ein weiterer Begriff, der beim räumlichen Hören und zur Beschreibung der Geräuschlokalisation genutzt wird, sind die Blauertschen Bänder. Diese wirken in der Medianebene beziehungsweise auf den Elevationswinkel ϑ , den man bereits aus Abbildung 1 entnehmen konnte. Anhand der Blauertschen Bänder wird erkannt, ob ein Geräusch von vorn, oben oder hinten kommt. Es wurde festgestellt, dass bestimmte Frequenzen einer bestimmten bevorzugten Richtung zugeordnet werden können. So werden zum Beispiel Geräusche, bei denen Frequenzen im Bereich von 1000 Hz einen erhöhten Pegel aufweisen, vom Ohr als von hinten kommend gedeutet. Im Gegensatz dazu werden Frequenzen im Bereich von 3 bis 4 kHz als von vorn kommend erkannt. Es gibt mehrere dieser Frequenzbänder für eine Richtung, die über das gesamte hörbare Spektrum verteilt sind. Dieses Phänomen basiert nur auf der Frequenzanalyse und ist demzufolge auch monaural, also mit nur einem Ohr, erkennbar.

In Abbildung 5 ist diese Richtungslokalisation der Medianebene noch einmal für den hörbaren Frequenzbereich grafisch dargestellt. [Blauert, 1997]

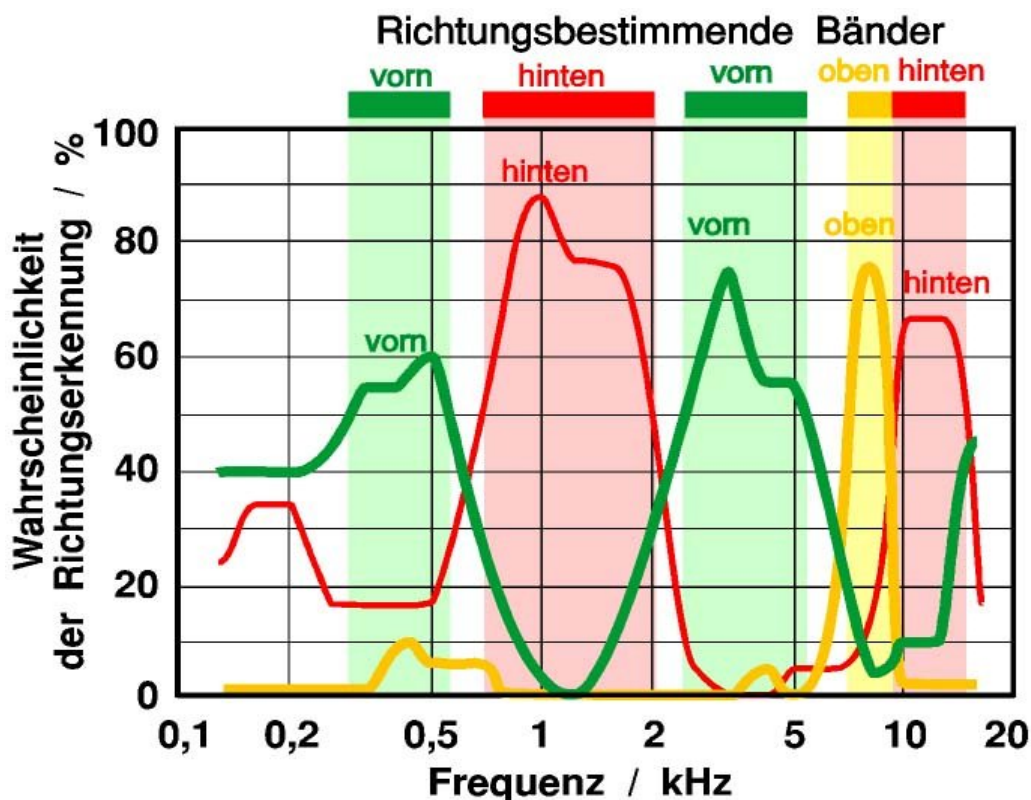


Abbildung 5: Richtungslokalisierung in der Medianebene [Hübner, 2011]

2.2 Head-Related Transfer Function

Eine binaurale Audio-Wiedergabe zielt darauf ab, einem Hörer den selben auditiven Eindruck zu vermitteln als wäre er in der abgespielten akustischen Szene selbst. Mittels einer solchen Wiedergabe, die auf Head-Related Transfer Functions (HRTF) zurückgreift, kann ein dreidimensionaler Eindruck der auditiven Szene über Kopfhörer erzeugt werden.

Durch die in dieser Arbeit thematisierte HRTF-Datenbank werden einige wichtige Signaleigenschaften, die für die dreidimensionale Wiedergabe benötigt werden, anhand eines Kunstkopfes aufgenommen.

Hierbei ist die Darstellung des Sounds über Kopfhörer noch nicht einmal begrenzt auf Im-Kopf-Lokalisation. Bei diesem über Kopfhörer häufig auftretenden Problem wird das Signal vom Hörer als direkt im Kopf befindlich angenommen.

Das Kunstkopfverfahren liefert bereits sehr realistische Ergebnisse. Jedoch gibt es auch hier Grenzen, weshalb man nun sehr häufig individualisierte HRTF's betrachtet, welche auch in einem späteren Kapitel behandelt werden. [Breebart, 2007]

2.2.1 Definition und Erklärung

Die head - related transferfunction ist in einer aufgenommenen Raumimpulsantwort (IR) enthalten, wobei die Impulsantwort für beide Ohren aufgenommen werden. Aus einer Impulsantwort können verschiedene Informationen für beliebige Räume herausgefiltert werden, wie zum Beispiel Absorptionsverhalten oder Nachhallzeiten.

Unter schalltoten Bedingungen beinhaltet die Impulsantwort nur den Einfluss des menschlichen Kopfes und könnte daher direkt nach der Aufnahme als HRTF zur binauralen Wiedergabe eingesetzt werden. Das aufgenommene IR - Paar stellt das Signalverhalten am Ohreingang im Zeitbereich dar und wird als Head - Related Impuls Response (HRIR) bezeichnet. Die HRTF ist nun die Fouriertransformierte dieses Signals (), also die Darstellung im Frequenzbereich,

$$HRTF_{gemessen, v}(\omega) = F\{HRIR_{gemessen, v}(t)\} \quad , v \in \{\text{links, rechts}\} \quad [2]$$

$F\{ \}$ stellt dabei die Fouriertransformation der in den Klammern enthaltenen Werte dar. Ein schallarmer Raum (eng.: anechoic Chamber) verzerrt allerdings dennoch ein wenig das Signal, da komplett schalltote Räume nicht existieren (man spricht daher eher von schallarmen oder reflexionsarmen Räumen). Aus dem Grund setzt man häufig das fouriertransformierte Signal an den Ohren in Bezug zur fouriertransformierten Raumimpulsantwort, welche separat in der Mitte des Kunstkopfes aufgenommen wurde [Zotkin, 2003]. Damit rechnet sich der geringe restliche Einfluss des Raumes und Einfluss der Lautsprecher in der HRTF heraus und die Formel [2] erweitert sich auf

$$HRTF_{Freifeld, v}(\omega) = \frac{F\{HRIR_{gemessen, v}(t)\}}{F\{RIR(t)\}} \quad , v \in \{\text{links, rechts}\} \quad [3]$$

Wobei RIR für die Raumimpulsantwort des leeren Raumes steht. In dieser Arbeit wurde die HRTF auch durch die Raumimpulsantwort geteilt, um das

bestmögliche Ergebnis zu erzielen. Dazu wurde ein Messmikrofon an den Punkt gesetzt, wo die Mitte des Kopfes bei den Messungen war. Ohne Einfluss des Kopfes wurde nun die Raumimpulsantwort mit der selben Methode wie die HRTF aufgenommen [Oldenburg, 2009; Faller, 2001]. Es gibt zwei weitere Möglichkeiten, die HRTF darzustellen:

$$HRTF_2 = \frac{(\text{Schalldruck am Eingang des offenen Ohrkanals})}{(\text{Schalldruck am Referenzpunkt bei abwesendem Hörer})}$$

$$HRTF_3 = \frac{(\text{Schalldruck am Eingang des geblockten Ohrkanals})}{(\text{Schalldruck am Referenzpunkt bei abwesendem Hörer})}$$

$HRTF_2$ stellt das Verhalten am Ohreingang dar, was für spätere Wiedergabe durch Kopfhörer gut geeignet ist. $HRTF_3$ vernachlässigt den Einfluss des Ohrkanals vollständig. Zur Analyse und zu Forschungszwecken wurde in dieser Arbeit die erste Darstellungsart verwendet, da sie das Signal direkt am Trommelfell repräsentiert. [Lindau, 2006; Moller, 1996]

2.2.2 Eigenschaften

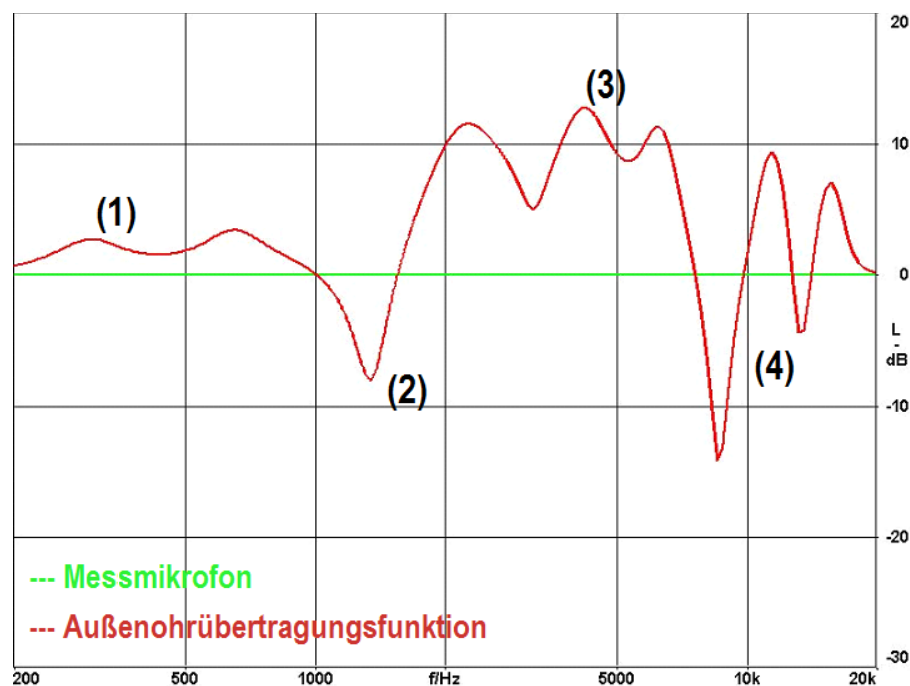


Abbildung 6: Grundform einer HRTF. [Fedke, 2007, S.5]

Beim Vergleich verschiedener HRTF's kann festgestellt werden, dass alle

auf einer gewissen Grundform in ihrem Aufbau basieren. Diese Grundform ist eng mit dem allgemeinen Aufbau des menschlichen Kopfes verbunden und wird davon beeinflusst. In der Abbildung 6 ist die Grundform anhand eines Beispiels HRTF aufskizziert. Im folgenden Abschnitt wird diese Grundform erläutert.

(1) In diesem Bereich wird das Verhalten der tieffrequenten Töne dargestellt. Da die Wellenlänge die Abmessungen des Kopfes übersteigt, sind hier kaum Verzerrungen zu verzeichnen.

(2) Beim ersten Einbruch der Übertragungsfunktion nehmen die Schultern des Menschen Einfluss auf die Übertragungsfunktion. Durch die Reflexionen an den Schultern kommt es zu einer destruktiven Interferenz bei etwa 1,4 kHz.

(3) Bei etwa 3 bis 4 kHz kommt es zu Überhöhungen aufgrund des Ohrkanals. Dieser hat mit seinen 2 cm Länge eine $\lambda/4$ -Resonanz bei 4 kHz zum tragen.

(4) Aufgrund der komplexen und feinen Struktur des Außenohrs kommen ab dem Bereich von über 3 kHz starke Interferenzen zum Tragen, die die HRTF gerade in diesen Bereichen sehr individuell gestaltet. [Fedke, 2007]

2.2.3 Binaurale Merkmale bei HRTF's

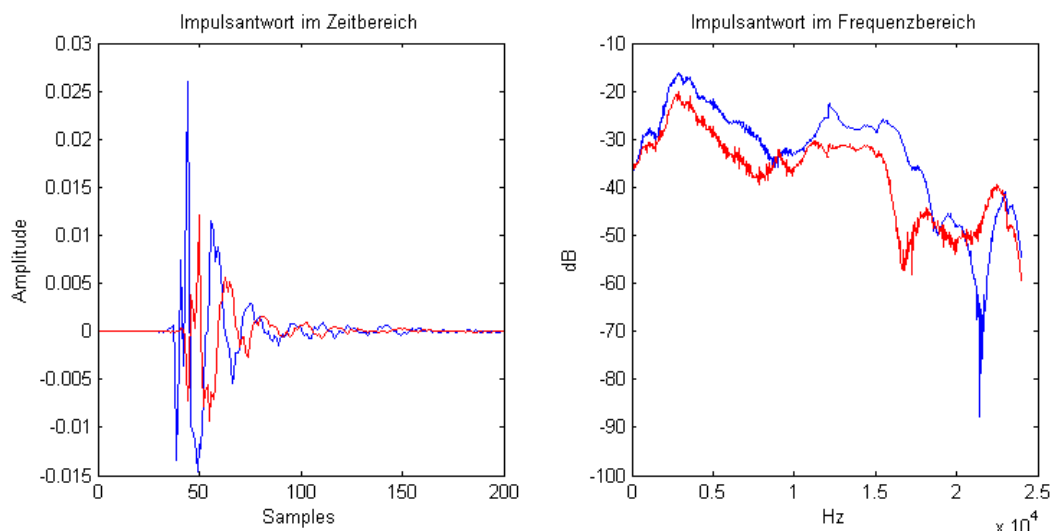


Abbildung 7: Paarbeispiele zweier HRIR (links) und HRTF (rechts), aufgenommen in einem schallarmen Raum. [Datenbank Oldenburg]

Anhand von den am Kunstkopf gemessenen HRTF kann man bestimmte Merkmale recht einfach erkennen. ILD und ITD sind sehr intuitiv durch

Vergleiche zwischen zwei HRTF - Paaren abzulesen. Man sieht in Abbildung 7, dass die Laufzeitdifferenz sehr gut im Zeitbereich (HRIR) durch die leichte Verschiebung der Kurven nachvollziehbar ist. Als Bezugspunkt wurde bei den Messungen in dieser Arbeit der Zeitpunkt genommen, wo der Anstieg mehr als 15 % des Maximums beträgt. Jedoch weist diese Variante bei scharfem seitlichen Schalleinfall, also Schalleinfallswinkel um die 90 ° und 180 °, sehr inkonsistente Werte auf, die zu schlechten Ergebnissen führen. Erkennbar ist dies an nicht starken Schwankungen in den ITD - Werten, was in Kapitel 4.1.2 näher erklärt wird.

Die Pegeldifferenz wird im Frequenzbereich anhand der Höhenverschiebungen der beiden Kurven abgelesen und ist, wie man leicht erkennen kann und auch schon in Kapitel 2.1.2 beschrieben wurde, frequenzabhängig.

2.2.4 Individualisierte HRTF

Obwohl binaurale Aufnahmeverfahren wie die Kunstkopfstereophonie sehr realistische Ergebnisse liefern, können starke Schwankungen zwischen einzelnen Personen auftreten. Gerade die Form des Außenohrs unterscheidet sich dramatisch zwischen einzelnen Personen oder auch zwischen Erwachsenen und Kinder und können auf die akustische Wahrnehmung negative Einflüsse ausüben, wie die falsche Bestimmung der Lokalisationsrichtung. Die wichtigsten Ursachen für diese Vielfalt sind Beugung, Reflexion und Streuung durch den Oberkörper, die Schultern und die Ohrmuscheln [Park, 2005]. Es können dabei auch unterschiedliche spektrale Verteilungen entstehen. Unterschiedliche kopfbezogene Übertragungsfunktionen sind die Folge, die bei jeder Person einzigartig sind und im allgemeinen als individualisierte HRTF's bezeichnet werden.

Die Abhängigkeit der Übertragungsfunktionen von individuellen Hörern ist ein Grund, warum weiterhin starker Forschungsbedarf besteht. Desweiteren lässt die akustische Reproduktion der binauralen Aufnahmen über Stereolautsprecher ein sehr unzufriedenes Klangbild erscheinen, da ein Übersprechen der Lautsprecher auf beide Ohren nicht zu vermeiden ist [Fedke,

2007]. Schnelle Methoden der Anpassung für einzelne Personen müssen noch entwickelt werden [**Zotkin, 2003**].

Um individualisierte HRTF's zu erstellen, sind verschiedene Methoden bekannt.

Die wohl einfachste wäre, um eine individualisierte HRTF für eine bestimmte Person zu erstellen, Messungen an dieser Person selbst durchzuführen, beziehungsweise einen Kunstkopf nach dessen Vorbild zu modellieren und diesen zu vermessen. Der enorme Aufwand, der dahinter steckt, wird schnell ersichtlich und lässt diese Methode als unwirtschaftlich herausfallen.

Eine weitere Methode ist die anthropometrische Messung. Aus einer Datenbank von HRTF's von mehreren Individuen wählt man die passendste anhand von Vergleichen in Bezug auf die Pinna, der medizinische Begriff für das Außenohr, aus. Verschiedene Größenwerte der Pinna der Testperson werden dabei mit denen der vorhandenen Datenbanken abgeglichen. Diese Variante ist wesentlich schneller als eine separate Kunstkopfmessung, aber auch vom Vorhandensein ausreichender Datenbanken abhängig. [**Cipic, 2001**]

2.2.5 Verwendung/Funktionsweise HRTF

Indem man die HRTF's eines Kunstkopfes bestimmt, ist man in der Lage, eine dreidimensionale Wiedergabe über Kopfhörer zu erzeugen. Es gibt dazu zwei grundlegende Möglichkeiten, dies zu realisieren.

Eine sehr einfache Möglichkeit der dreidimensionalen Wiedergabe ist, die Aufnahme direkt mit dem Kunstkopf durchzuführen. Die Aufnahme ist hierbei eine Stereoaufnahme mit Mikrofonen, die sich nahe des Ohreingangs des Kunstkopfes befinden. Wenn diese Aufnahme mit Kopfhörern wiedergegeben wird, entsteht ein realistisches dreidimensionales auditives Bild, welches den räumlichen Eindruck der Szene widerspiegelt. Es bleiben Fehlerquellen wie die fehlende Kopfhörerentzerrung und nicht individualisierte HRTF. Der Kunstkopf stellt hierbei die Position des Hörers dar. Eine prinzipielle Funktionsweise wird in Abbildung 8 aufgezeigt. Bei dieser Methode wird allerdings ein reflexionsarmer Raum benötigt.

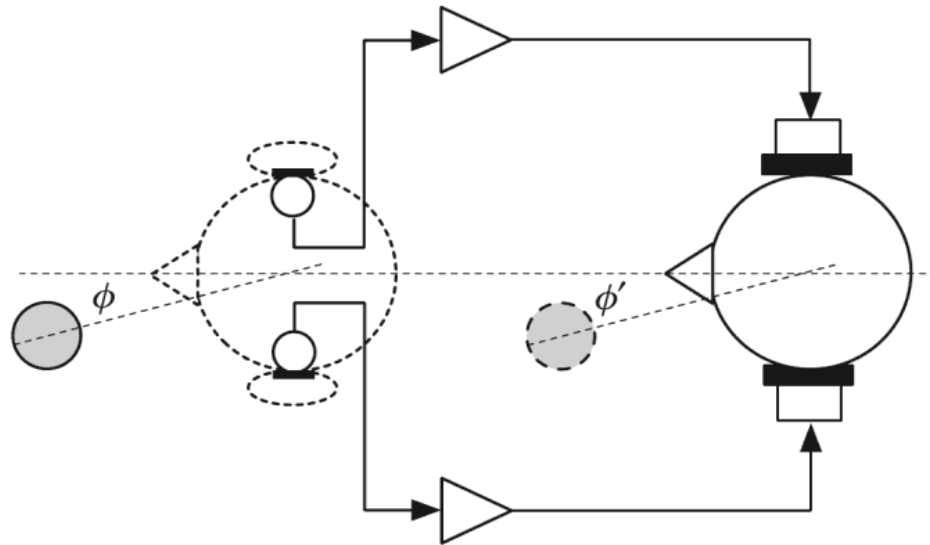


Abbildung 8: Binaurale Aufnahme mittels Mikrofonen im Kunstkopfhörorgan (links) welche über Kopfhörer wiedergegeben werden (rechts). [Breebart, 2007, S.10]

Binaurale Aufnahmen können auch erstellt werden, in dem man künstlich verschiedene Audiosignale zusammenmischt. Dazu wird ein Rendering System genutzt, welches von einer Quelle i ein Signal $x_i(t)$ im Zeitbereich entnimmt. Das Rendering zu einer bestimmten Position im Raum erfolgt, indem man das Signal mit einem HRIR - Paar, $h_{\text{links},i}(t)$ und $h_{\text{rechts},i}(t)$ faltet, welches die angenommene Position im Raum darstellt. Daraus entsteht ein binaurales Signal $y_{\text{links},i}(t)$ und $y_{\text{rechts},i}(t)$, dass den beiden Ohren zugeordnet werden kann. Die Faltung der beiden Signale ist mathematisch als

$$y_{(v,i)} = \int_{(\tau=0)}^{\infty} x_i(t-\tau) \cdot h_{(v,i)}(\tau) d\tau \quad v \in \{\text{links, rechts}\} \quad [4]$$

darzustellen. Das v steht für das jeweilige Ohr und i kennzeichnet das entsprechende Sample. Sie wird häufig aufgrund der einfacheren Rechnung und damit verbunden der besseren Rechengeschwindigkeit auch als Multiplikation im Frequenzbereich beschrieben:

$$Y_{(v,i)}(f) = X_i(f) \cdot H_{(v,i)}(f) \quad [5]$$

Wobei $H(f)$ der HRTF entspricht, $X(f)$ für das zu faltende Signal im Frequenzbereich und $Y(f)$ dem binauralen Signal im Frequenzbereich.

Die resultierenden Signalpaare der einzelnen Geräusche werden zusammenaddiert zu einem einzigen Signalpaar. [Breebart, 2007]

Bei dieser Methode ist die Genauigkeit auf diskrete Werte der aufgenommenen HRTF Datenbank begrenzt. Es ist jedoch möglich, mittels Interpolationsmethoden Zwischenwerte für vorhandene HRTF - Richtungen zu berechnen. Verschiedene Bestreben in der Hinsicht haben unter anderem zur spherical spline method [Carlile, 2000] und zur pole-zero approximation geführt [Blommer, 1997].

2.3 Stereophonieverfahren

Dieses Kapitel soll eine grobe Übersicht geben über Stereoaufnahmeverfahren, welche bisher zum Einsatz gekommen sind. Solche Verfahren benötigen zur Aufnahme immer zwei Mikrofone, um Stereo überhaupt erst zu ermöglichen. Die verschiedenen Methoden wurden im Laufe der Zeit immer weiter verbessert und vervollständigt. Es soll einen Einblick geben in die zeitliche Entwicklung der binauralen Aufnahmeverfahren und wie sich letztendlich die HRTF - Technik daraus entwickelt hat.

2.3.1 Differenzstereophonie

Eine der einfachsten und ältesten Varianten zur Stereoaufnahme ist die Differenzstereophonie, die sich wiederum aufteilt in Intensitäts- und Laufzeitstereophonie. Die Richtungsinformation für den Hörer wird hierbei nur aus den Laufzeitdifferenzen beziehungsweise den Pegeldifferenzen ermittelt. Das zeigt, dass die Methode sehr einfach und auch ungenau in Bezug auf die Räumlichkeit ist, da sie nur ein Merkmal verarbeiten kann. Vorteil dieses Verfahrens ist die gute Monokompatibilität. [Ederhof, 2004]

Der Aufbau erfolgt mittels zweier Druckempfängern, also Kugelmikrofonen ohne Richtcharakteristik, welche in einem Abstand zueinander aufgestellt werden für die Laufzeitdifferenz. Für die andere Variante werden zwei Nierenmikrofone in einem Winkelversatz übereinander (optimal ineinander) angebracht. Beide Varianten sind in Abbildung 9 und Abbildung 10 dargestellt.

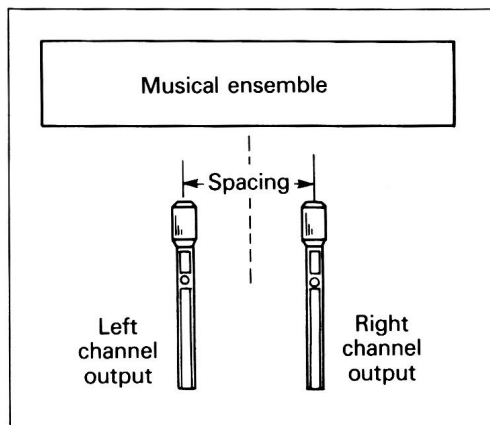


Abbildung 9: Mikrofonanordnung für Laufzeitstereophonie. [Bartlett, 1991, S. 19]

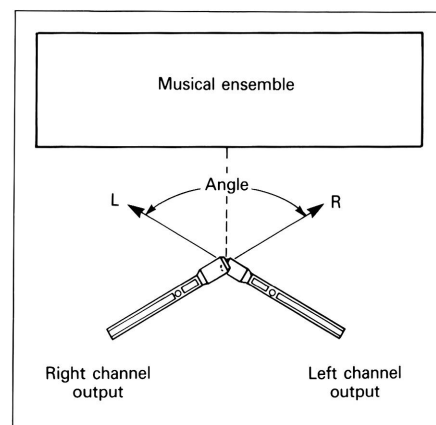


Abbildung 10: Mikrofonanordnung für Pegeldifferenzstereophonie. [Bartlett, 1991, S. 21]

2.3.2 Äquivalenzstereophonie

Die weiterentwickelte Version des Differenzstereophonieverfahrens ist die Äquivalenzstereophonie. Bei dieser werden die Richtungsinformationen sowohl aus der Laufzeitdifferenz als auch aus der Pegeldifferenz gewonnen.

Es bleibt bei zwei Mikrofonen, die verwendet werden. Jedoch werden diesmal Druckgradientenempfänger, genauer Nierenmikrofone mit Richtcharakteristik eingesetzt, welche in einem Winkel α zueinander verdreht werden. Der Abstand zwischen den Mikrofonen ist geringer als bei der Laufzeitstereophonie. Dieses Verfahren erzielt bereits einen guten Kompromiss zwischen Lokalisierbarkeit und Räumlichkeit des Sounds. Das bekannteste Beispiel dafür ist das ORTF - Verfahren, welches in Abbildung 11 skizziert ist. [Ederhof, 2004]

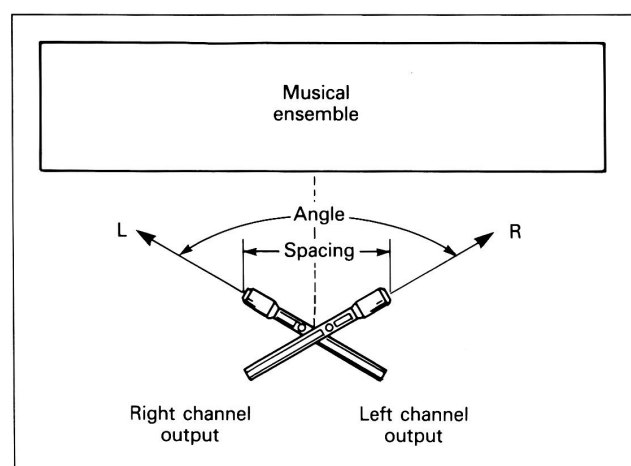


Abbildung 11: Grundsätzlicher Aufbau der Äquivalenzstereophonie, z.B. ORTF. [Bartlett, 1991, S. 23]

2.3.3 Trennkörperstereophonie

Eine weitere Anordnung nennt sich Trennkörperstereophonie. Bei dieser werden wieder Druckempfängermikrofone verwendet, die durch eine Trennscheibe, wie zum Beispiel der Jecklin - Scheibe, mit einem Durchmesser von 30 cm getrennt und im Abstand von 17 cm aufgestellt werden. Die Jecklin - Scheibe ist eine bekannte Anwendung und wird in Abbildung 12 aufskizziert.

Bei dieser Methode können Laufzeitdifferenz und auch Pegeldifferenz ermittelt werden, sowie einige Spektraldifferenzen, die durch die Abschattung der Jecklin-Scheibe auftreten. Nachteil ist, dass der seitliche Schalleinfall zu spitz verläuft, also die Beugung nur an einer scharfen Kante erzeugt wird und Kammfiltereffekte durch Reflexionen auftreten können. Allerdings erhält man eine gute Tiefbasswiedergabe und die Stereotrennung ist trotz des einfachen Aufbaus sehr gut. [Ederhof, 2004]

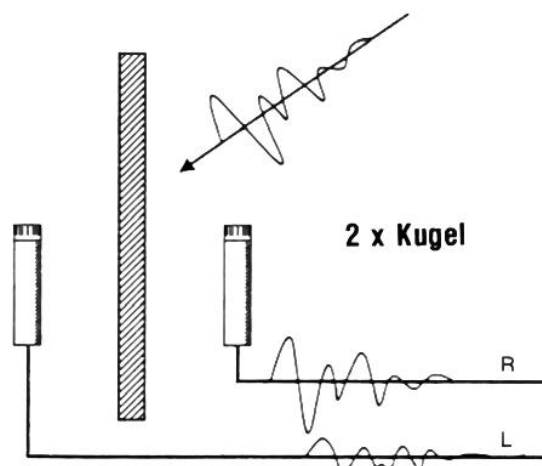


Abbildung 12: Aufbau einer Mikrofonierung mittels Jecklin-Scheibe mit seitlichem Schalleinfall.
[Curd, 2011]

2.3.4 Kugelflächenstereophonie

Um eine noch bessere Kopfhörerwiedergabe zu ermöglichen, hat man begonnen, sich am Kopf des Menschen zu orientieren und Stereophonieverfahren nach dessen Vorbild zu erschaffen. Daraus ist die Kugelflächenmikrofonie entstanden. Man erstellt eine Kugel mit einem Durchmesser von etwa 20 cm und bringt zwei Druckempfänger an beiden Seiten der Kugelfläche an. Dies könnte man schon als eine Simulation eines

Kopfes betrachten, wie man anhand von Abbildung 13 leicht erkennen kann.

Es werden bei dieser Stereophonieart die selben binauralen Merkmale extrahiert wie bei dem Trennscheibenverfahren.

Allerdings besteht hier der Vorteil, dass keine Kammfiltereffekte auftreten und der seitliche Schalleinfall nicht so spitz verläuft. Das Stereosignal bleibt dabei immer noch lautsprecherkompatibel. [Ederhof, 2004]



Abbildung 13: Beispiel eines Kugelflächenmikrofons. [Ederhof, 2004, S.212]

2.3.5 Kunstkopfstereophonie

Die nächste Stufe dieses Aufnahmesystems ist die Kunstkopfstereophonie. Eine recht realistische Nachbildung eines menschlichen Kopfes einschließlich der detailreichen Ohren ermöglicht einen sehr realistischen Klangeindruck. Dazu ist ein Beispiel in Abbildung 14 zu sehen. Ab dieser Stereophonieart kann man bei den Übertragungsfunktionen von kopfbezogenen Übertragungsfunktionen (HRTF) sprechen.

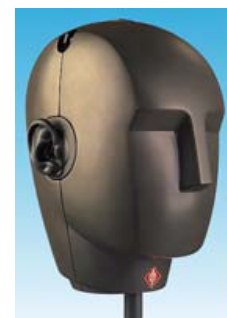


Abbildung 14: Beispiel eines Kunstkopfes von der Firma Neumann (KU100). [Fedke]

Sämtliche binaurale Merkmale, wie in Kapitel 2.1 beschrieben, können damit erfasst werden, einschließlich spektraler Differenzen, selbst der durch die Ohren verursachten. Jedoch ist diese Methode nur stark eingeschränkt lautsprecherkompatibel und fast ausschließlich bei Kopfhörerwiedergabe sinnvoll einsetzbar. Dies lässt sich auf das Übersprechen der Lautsprecher auf

jeweils beide Ohren zurückführen, was den Klang der beiden Signale verfälscht, da sie jeweils nur für ein Ohr bestimmt sind. [Ederhof, 2004]

2.3.6 Head and Torso Simulator

Die letzte Stufe der Optimierung von Kopfhörerwiedergabe erfolgt durch die Stereophonieaufnahme mittels eines Head-and-Torso-Simulators. Es wird ein Teil des Torsos des Menschen ebenso modelliert und mit dem Kunstkopf zusammen verbaut, wie man anhand von Abbildung 15 noch einmal nachvollziehen kann.

Damit kann man nun auch die letzten spektralen Beeinflussungen im mittelfrequenten Bereich betrachten, welche durch die Reflexionen an den Schultern zustande kommen. [Fedtke, 2007]



Abbildung 15: G.R.A.S. Kemar Head and Torso Simulator BA45. [verwendeter Kunstkopf]

2.4 Raumimpulsmessungen

Um Raumimpulsantworten messen zu können, muss der Raum zunächst mit einem geeigneten Signal angeregt werden. Verschiedene Methoden dieser Anregung werden in diesem Kapitel erläutert.

2.4.1 Klassische Messung

Die älteste und einfachste Methode, einen Raum anzuregen, ist mittels einer einfachen Impulsanregung. Dies könnte zum Beispiel ein Pistolenschuss oder ein platzender Luftballon sein. Oft wurde auch ein Knallfunkensender benutzt, um den Raum anzuregen. Da die spektrale Verteilung dieser Vorgänge bekannt ist, ist eine Reproduktion der Raumimpulsantwort möglich. Sie besitzen ein annähernd weißes Spektrum, also alle Frequenzen werden etwa gleich stark angeregt.

Eine weitere Möglichkeit ist die Durchführung mittels rosa Rauschen. Es ist ein zufälliges Rauschen, welches im tieffrequenten Bereich eine erhöhte Amplitude besitzt. Pro Oktave wird der Schallpegel um 3 dB verringert. Im Gegensatz zu weißem Rauschen, welches ein gleich verteiltes Spektrum besitzt, ist es wesentlich schonender für Geräte und für die messenden Personen, kann aber dafür genauso einfach generiert werden.

Bei diesen beiden Verfahren ist allerdings keine genaue Kenntnis über die Zeitfunktion vorhanden. Daher ist die spektrale Verteilung nur im Zeitmittel möglich und die Auswertung kann auch nur mit einer begrenzten Genauigkeit vollzogen werden. [Möser, 2009]

2.4.2 Sweep Verfahren

Bei dem Sweep - Verfahren handelt es sich um ein Gleitsinus mit ansteigender Frequenz. Dank steigender Rechenleistung ist er nun sinnvoll einsetzbar und ein sehr bedeutsames Messverfahren geworden. Durch die stetige Funktion der Frequenz ist es unanfällig gegenüber geringen Zeitinvarianzen. Auch Subjektiv sind hier schon sehr gut Störungen im Frequenzspektrum wahrnehmbar. Ein Sinussweep hat ein sehr markantes Erscheinungsbild im Zeitbereich, welches in Abbildung 16 zu erkennen ist.

Prinzipiell kann man hier wieder zwei mögliche Varianten unterscheiden. Zum einen den weißen Sweep mit konstantem Spektrum und den pinken Sweep (auch bekannt als Log - Sweep), dessen Spektrum bei höheren Frequenzen abfällt.

Der Fakt, dass der Sweep alle Frequenzbereiche nacheinander durchläuft, ist von Vorteil, wenn man die Impulsantwort eines Systems auf einen bestimmten Frequenzbereich reduzieren möchte. [Möser, 2009]

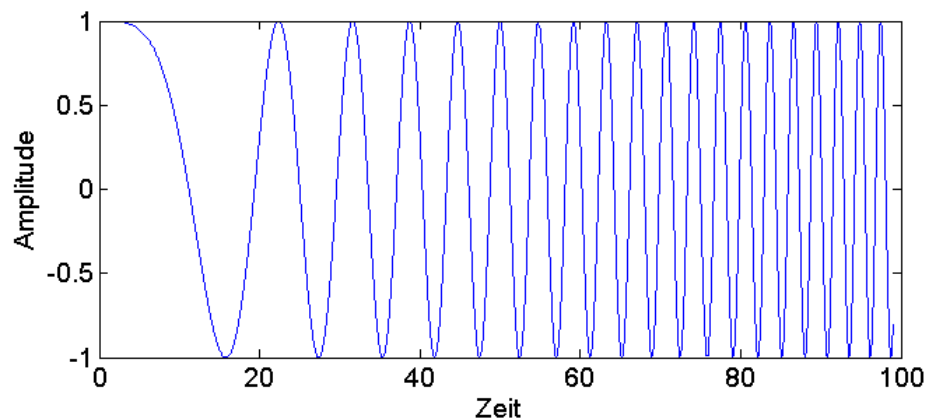


Abbildung 16: Beispielhafter Sinussweep im Zeitbereich. [eigene Simulation]

2.4.3 TDS-Verfahren

Time Delay Spectrometry oder kurz TDS ist eine Erweiterung des Sweep-Verfahrens. Neu dabei ist, dass eine Zweikanalmessung durchgeführt wird. Der Sinussweep wird in das System eingespeist, von dem die Systemantwort aufgenommen wird. Wird das Spektrum der Systemantwort durch das Spektrum des Originals dividiert, kann der Realteil der Impulsantwort errechnet werden. Dieselbe Division mit einem Kosinussweep anstelle des Sinussweep ergibt den Imaginärteil aufgrund der 90° Phasenverschiebung. Anschließend werden die Signale durch einen Bandpassfilter gefiltert, dessen Mittenfrequenz mit der Sendesignalfrequenz mitläuft.

Der Vorteil dieses Verfahrens ist ein sehr hoher Signal - Rauschabstand. Ebenso können effektiv Messungen auch in nicht - echofreier Umgebung durchgeführt werden. [Möser, 2009]

2.4.4 MLS-Technik

Die MLS-Technik (engl.: Maximum-length-sequence) ist eine neuere Technik, die auf der Konstruktion von Pseudozufallsfolgen basiert. Pseudozufall bedeutet hierbei, dass die Rauschfolge im Zeitbereich bekannt ist, da sie durch einen Konstruktionsalgorithmus erstellt wird. Die Zufallsfolge besteht nur aus binären Zuständen, also „0“ und „1“. Außerdem wird ihr eine ganzzahlig positive Ordnung N gegeben, welche zudem ihre Länge von $2^N - 1$ Samples bestimmt. Die Ordnung N muss an die Länge der Systemantwort angepasst werden, also wird für längere Nachhallzeiten auch eine höhere Ordnung gefordert. Ein Beispielrauschen befindet sich unter Abbildung 17. Das Spektrum des MLS - Rauschens ist Weiß bis auf den Gleichanteil.

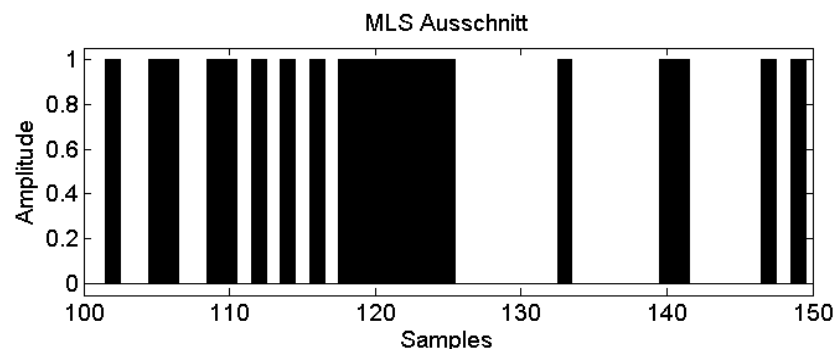


Abbildung 17: Ausschnitt eines MLS - Rauschens der Ordnung 8.
[eigene Simulation]

Es wird zur Erstellung des Rauschens ein Konstruktionsalgorithmus verwendet, der eine Exklusiv - Oder - Verknüpfung über ein Schieberegister erzeugt, diesen Wert rekursiv an das Register weiter gibt und zur Berechnung der nächsten Werte verwendet.

Durch die schnellen Wechsel der maximalen Auslenkungszustände kann es zu Verzerrungen und zum Übersteuern führen. Auch Nichtlinearitäten werden kaum toleriert. Die MLS - Technik hat allerdings auch spezielle Eigenschaften, die es ermöglichen, durch Korrelationsberechnung des MLS - Rauschens mit der Systemantwort des Raumes direkt die Impulsantwort zu erhalten. Somit hat die MLS - Technik historisch den Grundbaustein für die Auswertung im Zeitbereich gelegt. Weitere Vorteile sind die einfache Erstellung einer solchen Folge sowie der hohe erreichbare Signal - Rauschabstand. [Vanderkooy, 1989; Farina, 2000]

3. Durchführung

3.1 Verwendete Geräte

Die Durchführung der praktischen Arbeit fand im *Fraunhofer Institut für digitale Medientechnologie (IDMT)* in Ilmenau statt. Dort befindet sich ein reflexionsarmer Raum mit der Genauigkeitsklasse 1 (DIN EN ISO 3745). Seine Frequenzabdeckung reicht von 63 Hz bis zu 20 kHz und seine Rohbaumaße betragen 12 : 7,5 : 9,5 m³.

Es wurde ein Roboterarm namens *ELF* von *FourAudio* benutzt, um den Kunstkopf daran befestigen und drehen zu können. Der Lautsprecher hatte eine konstante Position während der Messungen. Das zu dem *ELF* gehörige Mess- und Aufnahmesystem *Monkey Forest* wurde benutzt, um Daten aufnehmen und speichern zu können. Zur Impulsanregung wurde ein logarithmischer Sinussweep vom Programm *Monkey Forest* verwendet, welcher über Lautsprecher von *Genelec* ausgegeben wird. [Four, 2008]

Der verwendete Kunstkopf ist ein *G.R.A.S. Kemar Head and Torso Simulator Type 45BA* mit einem Kopfdurchmesser von 17,5 cm. [Kemar, 2008]

Zu Beginn traten noch einige Schwierigkeiten auf. Zum Beispiel ragte der Kunstkopf bei konventioneller Befestigung über den Drehmittelpunkt des Roboterarmes von *FourAudio* hinaus. Wenn dies der Fall ist, kommt der Schall nicht mehr mit dem gewünschten Winkel beim Ohr an. Es entsteht ein Winkerversatz, welcher die Messung verfälschen kann und im nach hinein nicht mehr zu entfernen ist. Um den Kopf dennoch effektiv befestigen zu können, musste eine neue Methode gefunden werden.

Dazu wurde entschieden, den *KEMAR* mit Bauschaum zu füllen, um eine Halterungsstange darin anbringen zu können und somit die Befestigung am *ELF* zu gewährleisten. Damit konnte der Kunstkopf über den Halterungsstab des Roboterarmes geschoben werden und der Versatz zum Drehmittelpunkt wurde beseitigt.

Dieses Vorhaben wurde allerdings verworfen, da es sich nicht als praktikabel erwiesen hat. An einer universelleren Lösung für spätere Messungen wird noch gearbeitet.

3.2 Messaufbau

Nachdem sehr lange Verzögerungen in Bezug zur Fertigstellung der Halterung für den Kunstkopf in Kauf genommen werden mussten, wurde zwangsweise ein Ersatzweg eingeschlagen. Dabei wurde entschieden, nur eine horizontale Ebene des Kunstkopfes zu vermessen, diese aber dafür mit den zeitnah vermessenen Probanden - HRTF's zu vergleichen. Damit musste also die Möglichkeit der kompletten Datenbank des Kunstkopfes verworfen werden. Allerdings erhält man so Daten von verschiedenen Kopfformen und es besteht nun die Möglichkeit, den Kunstkopf mit realen Personen zu vergleichen. So kann man entscheiden, ob die verwendeten Methoden zur Erstellung der ITD und ILD für reale Personen genauso zuverlässig sind, wie für den Kunstkopf oder welche Unterschiede im Zeit - und Frequenzbereich zu Problemen bei Probandenmessungen führen können. Die Auswertung wurde in Kapitel 4.2.3 durchgeführt.

Zu dem Zweck werden vier Lautsprecher 90° versetzt zum Kunstkopf in einem Abstand von etwa 130 cm aufgestellt. Die Probanden ebenso wie der Kunstkopf werden auf einem Sitz positioniert, der anstelle des Roboterarmes auf dessen Grundhalterung befestigt wurde. Durch die vier Lautsprecher muss man die Probanden nur um 90° drehen, da an jeder Position vier HRTF - Paare aus um jeweils 90° versetzten Richtungen aufgenommen werden können. Der grundsätzliche Messaufbau findet sich in Abbildung 18 wieder, sowohl als Foto als auch als Prinzipskizze. Der Kunstkopf wurde mit einer 1° genauen Auflösung in horizontaler Richtung vermessen. Ein Laser dient der Kalibrierung der korrekten Positionierung des Kunstkopfes auf dem Stuhl und der Lautsprecherpositionen. Der Laser ist in Abbildung 19 dargestellt.

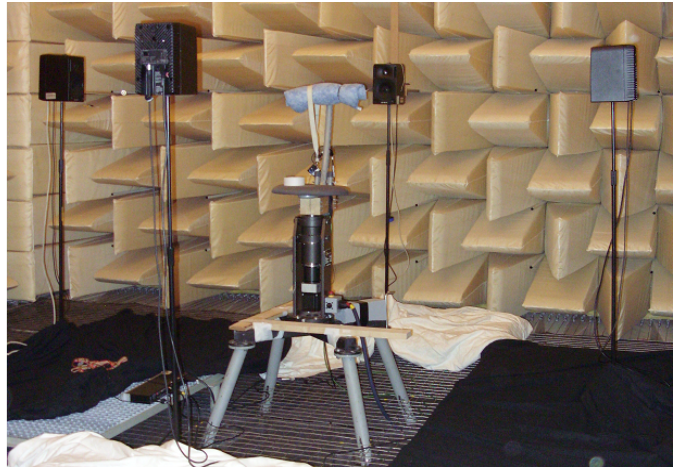
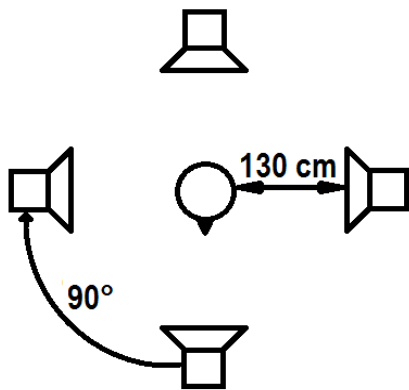


Abbildung 18: Grundsätzlicher Messaufbau im Schallarmen Raum, links: Prinzipskizze, rechts: Fotodarstellung



Abbildung 19: Laser zur korrekten Messpositionierung.

4. Auswertung

Nach Durchführung der Messung und Aufnahme der kopfbezogenen Impulsantworten wurden noch einige Nachbearbeitungen der erhaltenen Messdaten vorgenommen. Die Nachbearbeitungen sind zusammen mit der Auswertung in diesem Kapitel zu finden.

4.1 Aufbereitung der Messergebnisse

4.1.1 Fensterfunktion

Nachdem die ersten Ergebnisse der Messung vorliegen, können die in den Grundlagen erläuterten Informationen extrahiert werden. In diesem Abschnitt werden die dazu notwendigen zusätzlichen Schritte beschrieben.

Anhand der Abbildung 20 sollen verschiedene Eigenschaften einer der gemessenen Impulsantworten erläutert werden. Die beiden folgenden Formeln zur Berechnung von Laufzeiten und Verzögerungen sind hierbei von Bedeutung: $c = \frac{s}{t}$ [6], $vz = t \cdot Fs$ [7].

Dabei ist vz die Verzögerung in Samples. Die Sample Rate Fs beträgt 48 kHz. Damit sollen einige Überlegungen bezüglich einer Fensterfunktion und der Verzögerungen einer Impulsantwort im Allgemeinen getroffen werden.

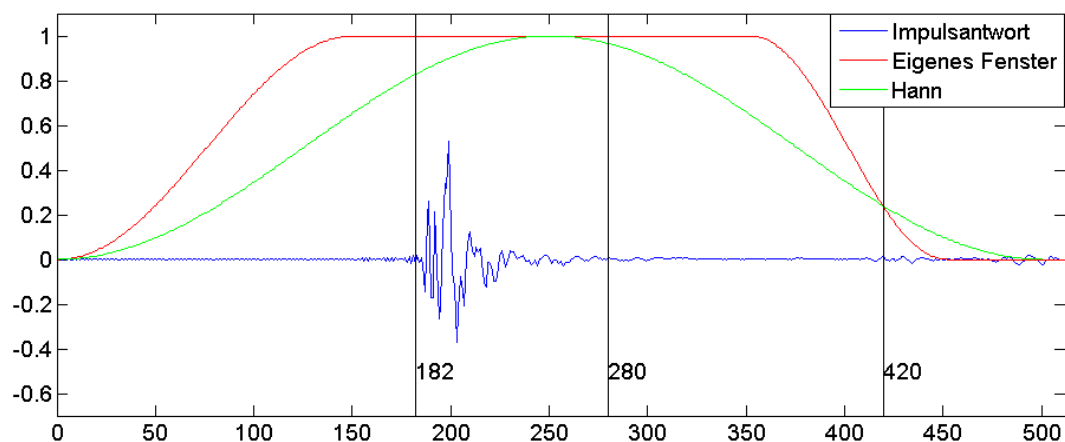


Abbildung 20: Beispielimpulsantwort der gemessenen Datenbank und passende Fensterung

- 1) Der Abstand vom Lautsprecher zum Kopf beträgt 130 cm. Aus den gegebenen Formeln ergibt sich eine Verzögerung v_z von 182 Samples. Die Abbildung lässt sehr gut erkennen, dass der Impuls knapp nach 182 Samples eintrifft. Da der Lautsprecherabstand sich während der Messung nicht ändert, kann sich auch die Anfangsverzögerung nicht ändern. So kann man durch diese Rechnung sicher sein, dass auch die anderen Impulsantworten der Datenbank eine Verzögerung in dieser Größenordnung aufweisen.
- 2) Erste Reflexionen, die über den Boden den Kunstkopf erreichen, können mit etwa 2 m Abstand zum Lautsprecher angenommen werden. Durch die beiden Formeln lässt sich auch hier die ungefähre Verzögerung von 280 Samples errechnen. Da der Boden mit schalldämmendem Material abgedeckt wurde und dieser Bereich noch in das Ausschwingen der eigentlichen Impulsantwort hineinfällt, wurde dieser Einfluss vernachlässigt.
- 3) Der letzte markierte Teil beschreibt das Eintreffen der ersten Reflexionen, die vermutlich von den Lautsprechern selbst kommen, da der Umweg von einem Lautsprecher über einen anderen hin zum Kunstkopf mit etwa 3 m angenommen werden kann. Die Berechnung durch Formel [6] und [7] ergibt eine Verzögerung von 420 Samples. Man erkennt auch deutlich, dass die Amplitude kurz vorher schon einiges geringer war, was den Verdacht bestätigt, dass diese Reflexionen nicht mehr zur benötigten Impulsantwort gehören.

Aus den drei genannten Punkten kann man nun eine Fensterfunktion erstellen. Eine Möglichkeit wäre, ein Hann - Window zu benutzen. In dem Fall würde aber, wie auch anhand der Abbildung erkennbar ist, der wichtige Teil der Impulsantwort ebenfalls ein wenig gedämpft. Die Dämpfung fällt zwar kaum ins Gewicht, wenn man den Frequenzgang vergleicht, sollte aber dennoch aus Gründen der Genauigkeit vermieden werden. Gerade ein zu kurz gewähltes Fenster kann sehr schnell zu einem Kammfiltereffekt und anderen frequenzdämpfenden Effekten führen. So wurde ein selbst erstelltes Fenster gewählt, welches einen Raised - Cosine als Ein - und Ausschwingfunktion benutzt. In der Mitte wird das Signal mit eins multipliziert. Nach dem Fenster

werden noch Nullen aufgefüllt, um eine gebräuchliche Länge der Impulsantwort von 512 Samples zu erzeugen. Der Rest der Impulsantwort wird demnach komplett abgeschnitten.

4.1.2 Differenzextraktion

Um wichtige Daten schnell zur Hand zu haben und diese ebenso schnell auswerten zu können, wurden wichtige binaurale Merkmale aus den HRTF herausgerechnet. Bei den herausgerechneten Werten handelt es sich um die in den Grundlagen behandelten interauralen Laufzeitdifferenzen und den interauralen Pegeldifferenzen, für deren Berechnung eine Matlab - Funktion geschrieben wurde.

Bei den Laufzeitdifferenzen wurde die Anstiegsvariante der Maximalvariante vorgezogen. Bei der Maximalmethode wird die Differenz anhand der Verschiebung zwischen den Maximalwerten der beiden Impulsantworten errechnet. Sehr ähnliche Ergebnisse liefert eine Methode, die über Korrelation versucht, die Laufzeitdifferenz zu bestimmen. Über die Anstiegsvariante wird ein Grenzwert des Anstiegs benachbarter Samples als Vergleichspunkt herangezogen. Die Anstiegsvariante ist wesentlich stabiler und liefert konsistente Werte, was man auch beim Betrachten der Laufzeitdifferenzen einer kompletten Ebene erkennen kann. Die Werte verlaufen also fast linear an - oder absteigend, wie man es für 90 ° Abschnitte erwarten würde. Eine Darstellung dazu befindet sich unter Abbildung 21. Die Anstiegsvariante bietet den Vorteil, veränderbare Werte als Anstiegssuche zu nutzen. Die Werte fallen meist sehr gering aus, womit rechtzeitig Überhöhungen detektiert werden. Problematisch wird es, wenn der Grenzwert des Anstiegs bei dieser Methode nur knapp erreicht oder verfehlt wird und somit der benötigte Anstieg nicht erkannt werden kann. Die Maximalmethode liefert bei kleinen Winkeln ebenfalls brauchbare Werte und bietet eine gute Alternative zur Anstiegsmethode. Gerade bei grenzwertigen Winkeln allerdings sind oft mehrere ähnlich hohe Maxima zu finden, welches die korrekte Auswahl behindert. Weil bei den in dieser Arbeit verwendeten Probandenmessdaten häufig mehrere Maxima auftreten, wurde sich für die Anstiegsvariante

entschieden.

Ebenso sind die errechneten $800\ \mu\text{s}$ aus Kapitel 2.1.1 als Sampledifferenzen wiederzufinden. Diese $800\ \mu\text{s}$ entsprechen 38 Samples bei einer Abtastrate von $48000\ \text{Hz}$ und das entspricht genau dem Maximum in der gegebenen Abbildung. Der Grenzwert für den Anstieg wurde auf 15 % festgesetzt, da dies die grafisch besten Ergebnisse geliefert hat.

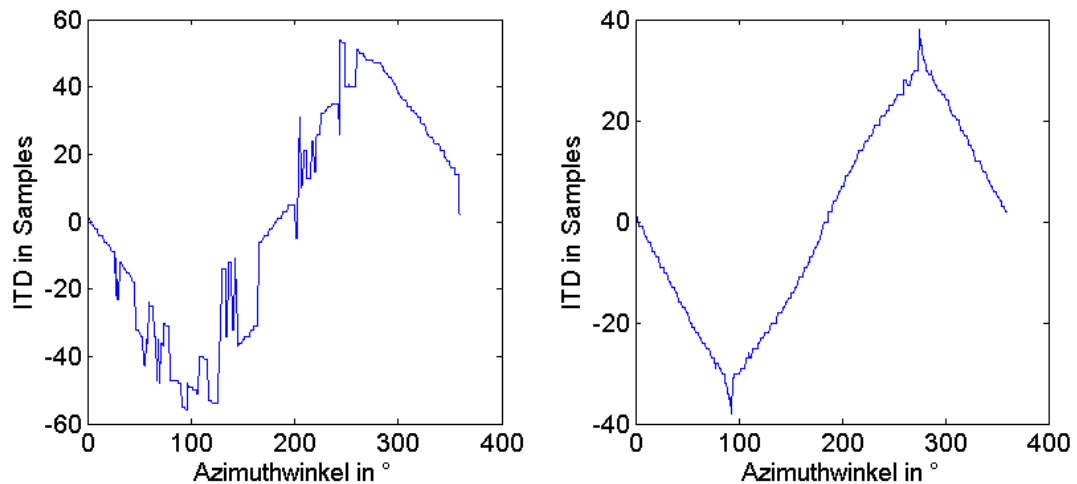


Abbildung 21: Vergleich des Ergebnisses beider Methoden zur ITD-Berechnung (links: Maximalwertsuche, rechts: Anstiegssuche (15 %)). [eigene Messung]

Da die interaurale Pegeldifferenz, wie in Kapitel 2.2.2 erklärt, frequenzabhängig ist, wurde dafür ein Vektor aus den Werten berechnet, der diese Differenz für alle abgedeckten Frequenzen darstellt.

Beide Differenzwerte wurden in die jeweiligen Ordner als ergänzende Daten gespeichert, sowie die Sampleverzögerungen zur Berechnung der Laufzeitdifferenz.

4.1.3 Minimalphasen Approximation

Nachdem die beiden genannten Eigenschaften der HRTF als separate Werte vorliegen, wird eine Minimalphasenapproximation durchgeführt. Dies führt zum Entfernen des zeitlichen Versatzes der Impulsantworten. Der Vorteil darin liegt, die Gruppenlaufzeiten sowie die Energieverzögerung minimal zu halten. Desweiteren kann die Datenmenge drastisch reduziert werden. Zusammen mit der Separierung kann das entstehende Amplitudenspektrum mit individuellen Laufzeitdifferenzen verknüpft werden. Da dieser Vorlauf nicht

einfach abgeschnitten werden kann, wird eine Minimalphasenreduktion vorgenommen.

Eine Minimalphasensequenz ist ein System, dessen Fourier - Transformierte minimalphasig ist, mit der Eigenschaft, dass das logarithmische Amplitudenspektrum und der Phasengang Hilberttransformationen voneinander sind. Somit wird auch die Berechnung eines solchen Systems ersichtlich: Man führt eine Hilberttransformation mit dem fouriertransformierten Signal durch, um so den neuen korrigierten Phasengang zu erhalten. Das nun entstandene System behält die ursprüngliche Übertragungsfunktion im Frequenzbereich (Amplitudenspektrum) bei. Die Funktion im Zeitbereich wird dadurch soweit nach vorn geschoben, bis der Gleichanteil eine Phasenverschiebung von 0° ausmacht. Ein Minimalphasensystem hat die geringste Energieverzögerung von allen realisierbaren Systemen. Die Gruppenlaufzeit ist ebenfalls minimal. Man sieht sehr deutlich in Abbildung 22, dass der „Schwerpunkt“ der Zeitfunktion sehr weit an den Koordinatenursprung gewandert ist. Auch zu beachten ist, dass die beiden Zeitfunktionen einander nicht mehr ähneln, also keine Verschiebung des Originals stattgefunden hat. Der Frequenzgang ist dennoch annähernd der gleiche für beide Funktionen und Unterschiede sind für die Wahrnehmung des

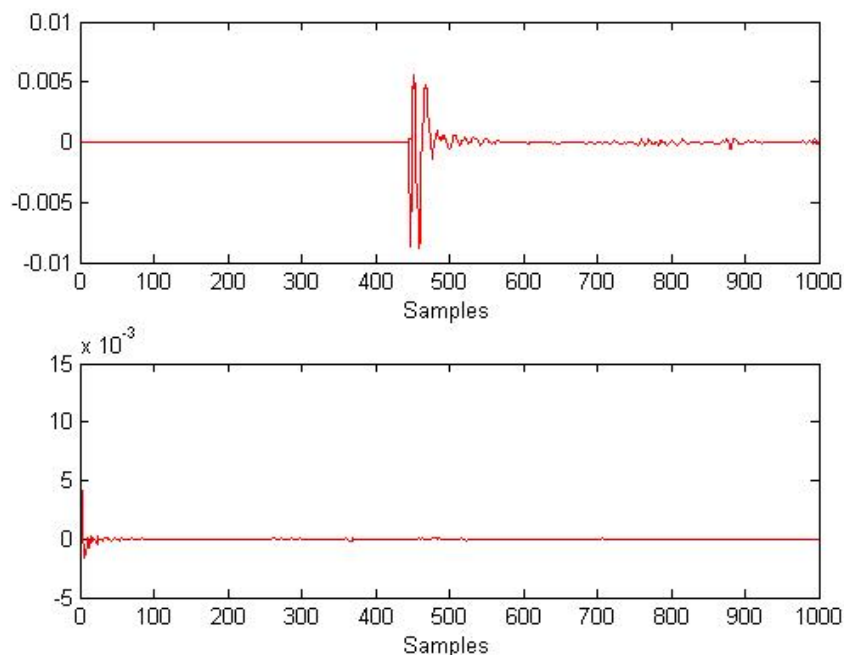


Abbildung 22: Oben: Impulsantwort eines Systems. Unten: Minimalphasenreduzierte Impulsantwort. [Datenbank Oldenburg]

Menschen so gering, dass sie oftmals vernachlässigt werden. Der Phasengang wurde durch die Hilberttransformation verändert. Er ist nun für den Frequenzgang einzigartig. [Kulkarni, 2002]

Leider sind bei diesem wie auch bei jedem anderen signalbearbeitenden Verfahren Fehlerquellen, die das Signal verfälschen, nicht auszuschließen. Es wird demzufolge noch eine Fehleranalyse durchgeführt, welche erkennen lässt, welche Fehler bei der Übertragungsfunktion durch die Minimalphasenreduktion entstehen und ob die verwendete Funktion zur Berechnung der Minimalphasenreduktion funktioniert und dazu geeignet ist. Eine Grafik zur grundlegenden Herangehensweise dazu findet sich unter Abbildung 23.

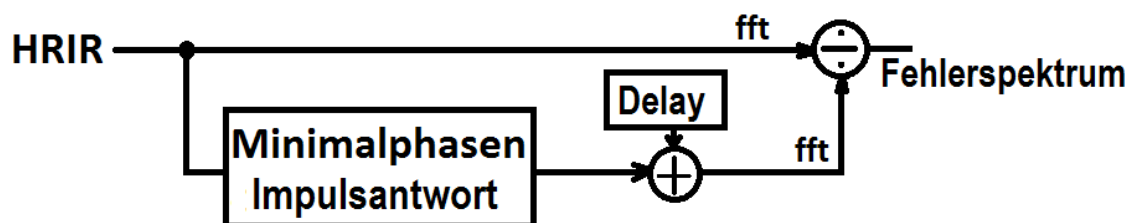


Abbildung 23: Abfolge der Signalverarbeitung.

4.2 Auswertung der Messergebnisse

4.2.1 Minimalphasenreduktion

Die Erstellung der Minimalphasenvariante der Impulsantworten wurde durch eine eigenständige Matlabfunktion durchgeführt. Wie schon in Kapitel 4.1.2. erklärt, ist der Unterschied zwischen Originaler Impulsantwort und der minimalphasenreduzierten Variante im Frequenzbereich sehr gering. Es wurde eine Überprüfung mit einer Fehlerbetrachtung der beiden Frequenzbereiche durchgeführt. In Abbildung 24 erkennt man sehr gut, dass kaum Fehler von 3 dB oder darüber auftreten. Allerdings sind vereinzelt gerade bei hohen Frequenzen Differenzen von bis zu 20 dB zu finden. Der größte Bereich liegt bei 0 dB Differenz, was zeigt, dass die Funktion zur Minimalphasenreduktion gut funktioniert.

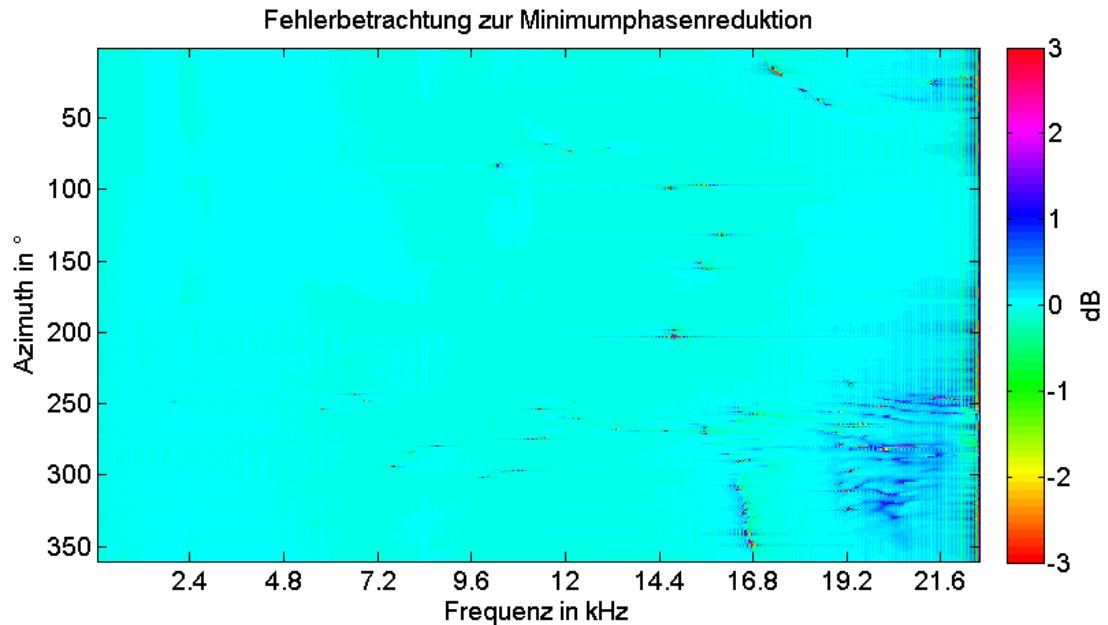


Abbildung 24: Betrachtung des Fehlers im Frequenzbereich der Minimalphasendaten in Bezug zum Original.

Es wurden die Originalen Daten mit einer rekonstruierten Minimalphasenapproximation verglichen. Dazu wurde der Minimalphasenvariante die zur Berechnung der Laufzeitdifferenz verwendeten Verzögerungswerte vorgestellt. Somit beginnt der Schwerpunkt wieder an der des Originals entsprechenden Stelle, wie sich anhand von Abbildung 25 erkennen lässt. Da die Phase aller Frequenzen minimiert wurde, ist der Bereich des Maximums des rekonstruierten Signals wesentlich kürzer und schwingt schneller aus als beim Original.

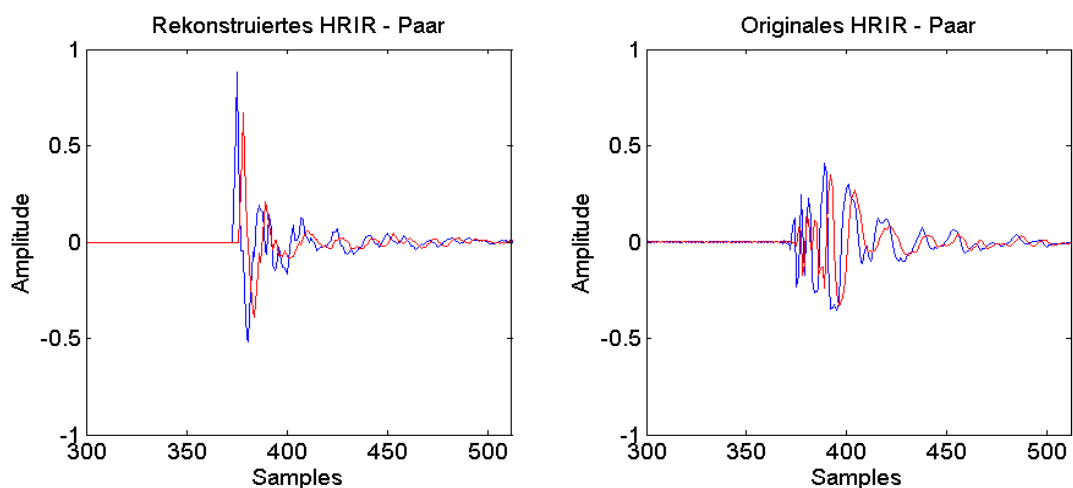


Abbildung 25: Beispiel eines rekonstruierten und eines Originalen HRTF - Paares (10° Azimut)

Eine Darstellung zum Vergleich der Phase und des Frequenzspektrums findet sich unter Abbildung 26. Es ist zu sehen, dass die Phase nach der

Rekonstruktion wieder einen ähnlichen Verlauf wie die des Originals aufweist.

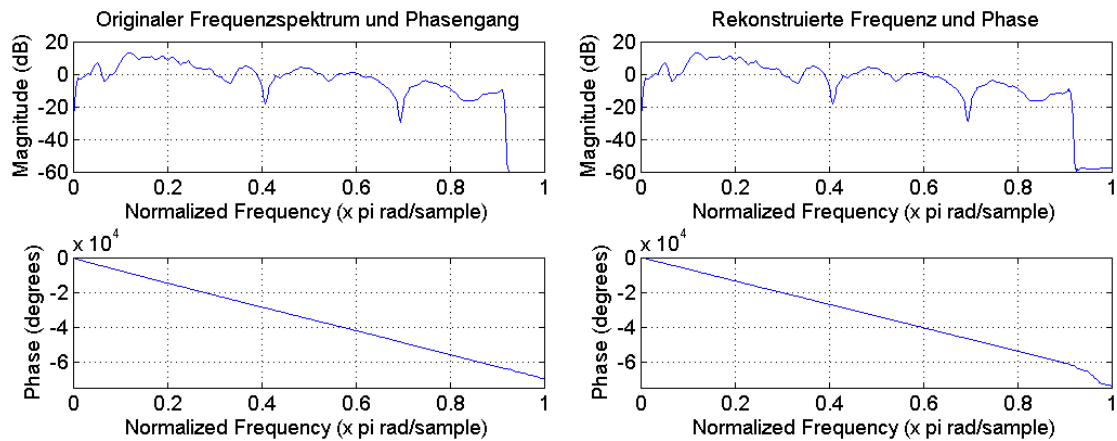


Abbildung 26: Vergleich der Phase und des Amplitudenspektrums von rekonstruierten und originalen Daten. (linkes Ohr, 10° Azimut)

4.2.2 Amplitudenspektrum

Die komplette Amplitudendifferenz (Abbildung 27) sowie die komplette Laufzeitdifferenz (Abbildung 21) wurde für alle Winkel geplottet. Ein sehr starker Abfall der Amplitude bei sehr hohen Frequenzen, ab etwa 22 kHz, deutet darauf hin, dass die Lautsprecher diesen Bereich nicht mehr abdecken. Der nicht benötigte Teil wurde nicht geplottet.

Es ist erkennbar, dass selbst der Kunstkopf keine korrekte Symmetrie mit der Symmetrieachse horizontal bei 180° aufweist.

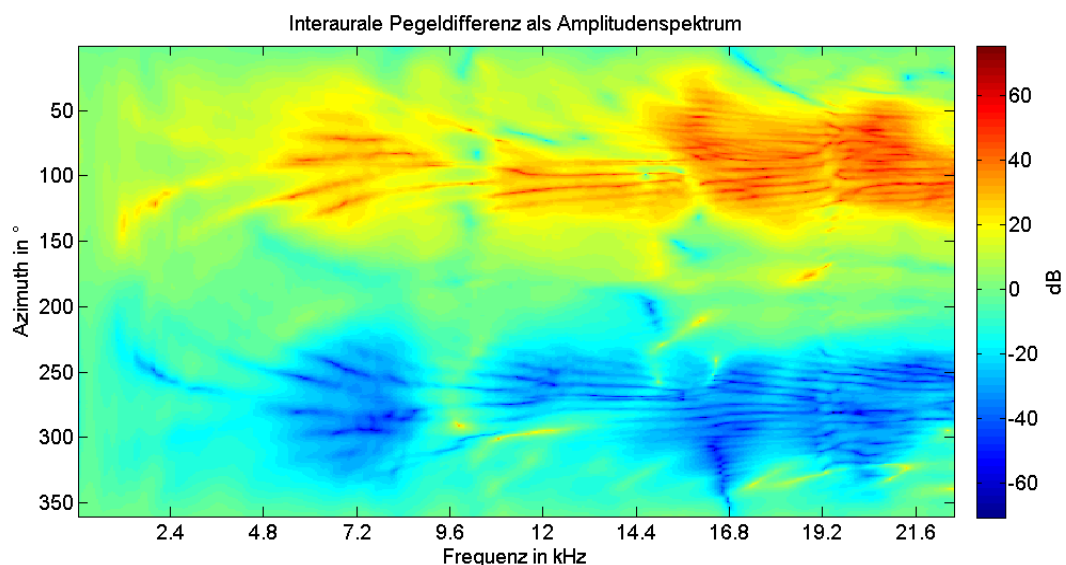


Abbildung 27: Amplitudenspektrum der Differenz beider Ohren als Farbintensitätsdarstellung

4.2.3 Probandenmessungen

Die Probandenmessungen wurden auf dieselbe Weise durchgeführt wie mit dem Kunstkopf. Es soll in diesem Kapitel anhand von zwei ausgewählten Probanden - HRTF's ein Vergleich zu dem Kunstkopf, der als Referenz dient, herangezogen werden. Die genutzten Datenbanken sind *beer.mat* und *ltu.mat*.

Als erstes sollen die erstellten ITD's verglichen werden. Der Kunstkopf als Referenz weist einen linearen Anstieg auf, der von der ITD erwartet wird. Es wird mit der Beer - Datenbank verglichen. Der Vergleich kann anhand von Abbildung 28 nachvollzogen werden.

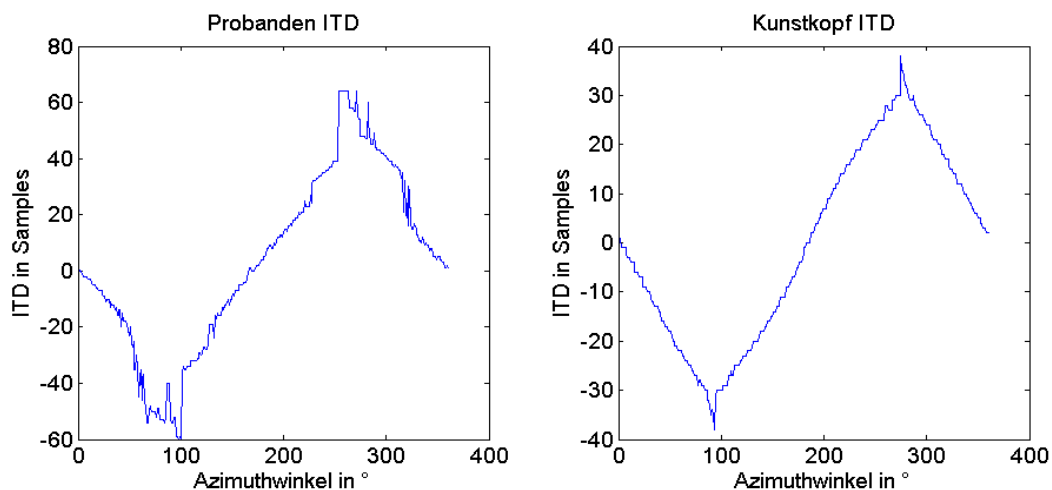


Abbildung 28: Laufzeitvergleich zwischen Proband (*beer.mat*) und Kunstkopf

Es ist erkennbar, dass die für den Probanden errechneten Werte weniger linear sind und deutlich andere Wertebereiche aufweisen. Dieser Formunterschied ist auf allgemeine Durchführungsfehler, die am Ende des Kapitels zusammengefasst werden, zurückzuführen. Die unterschiedliche Kopfform des Probanden zum Kunstkopf sollte nur die Höhe des ITD - Versatzes beeinflussen. Einige sehr starke Sprünge sind in der ITD des Probanden gerade bei grenzwertigen Winkeln um 90° und 270° zu sehen. Sehr auffällig ist der Sprung von 253 ° zu 254 °.

Um diesen Sprung näher zu betrachten, wird als nächstes die Zeitfunktion der HRTF's anhand von Abbildung 29 verglichen.

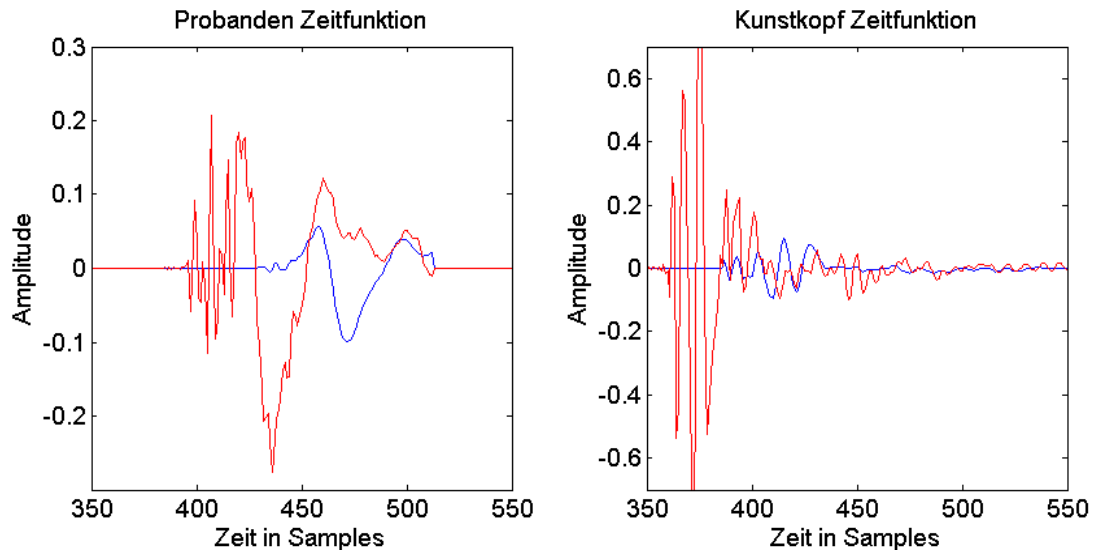


Abbildung 29: Vergleich der Zeitfunktion von Proband (*beer.mat*) und Kunstkopf für einen Azimut von 254°

Es wird erkennbar, dass die Zeitfunktion des Probanden individuell gestaltet ist und der Ausschwingvorgang sogar mit einer Überhöhung zusammenfällt. Im Gegensatz dazu kommt beim Kunstkopf direkt nach den Überhöhungen der Ausschwingvorgang. Für das genannte Problem bei der ITD - Berechnung muss das Signal des der Schallquelle abgewandten Ohres (in der Grafik blau dargestellt) betrachtet werden. Beim Probanden finden sich dagegen nur wenige langgezogene Überhöhungen. Aus der grafischen Auswertung lässt sich entnehmen, dass der reale Kopf einen etwas anderen Tiefpasscharakter aufweist als der Kunstkopf (Kammfilter - ähnlich). Ursachen dafür sind der individuelle Außenohraufbau jedes Menschen, der Innenohraufbau sowie die Haare der Probanden. Durch diese Tiefpassfunktion wird die Anstiegssuche erschwert, die den Grenzwert bei 430 Samples für 253° entdeckt hat. In der Darstellung von 254° wurde die Anstiegsschwelle allerdings erst nach 450 Samples detektiert. Somit muss die erste Schwelle bei 254° auf über 15 % gestiegen sein. Diese Sprünge können kaum vermieden werden, da selbst bei Änderung des Schwellwertes eben diese Probleme an anderer Stelle auftreten würden.

Zuletzt wird das Amplitudenspektrum verglichen. Ein Gesamtvergleich wird in Abbildung 30 vorgenommen. Dazu wurde die Datenbank *ltu.mat* herangezogen.

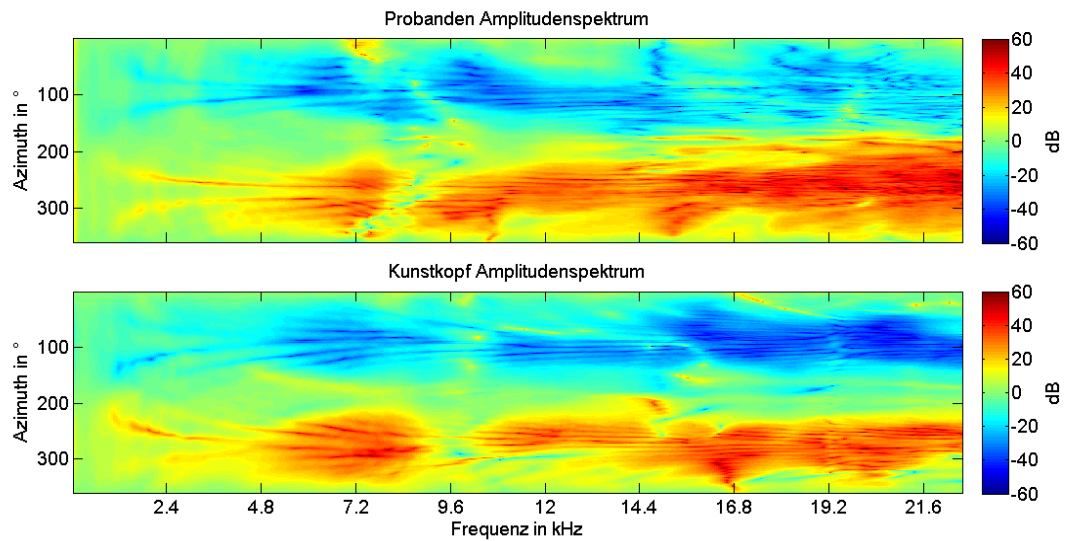


Abbildung 30: Amplitudendifferenzvergleich von Proband (ltu.mat) und Kunstkopf

Durch den Vergleich wird deutlich, dass die Grundformen einander stark ähneln. Ebenso stimmen die Wertebereiche überein. Eine Symmetrie mit der Symmetrieachse bei 180° ist bei beiden Plots erkennbar. Zum direkten analytischen Vergleich wurde die Differenz der beiden Grafiken in Abbildung 31 geplottet. Dabei ist erkennbar, dass trotz der gleichen Grundform teilweise erhebliche Pegelunterschiede auftreten. Das zeigt wiederum, dass individualisierte HRTF sehr Variantenreich sein können.

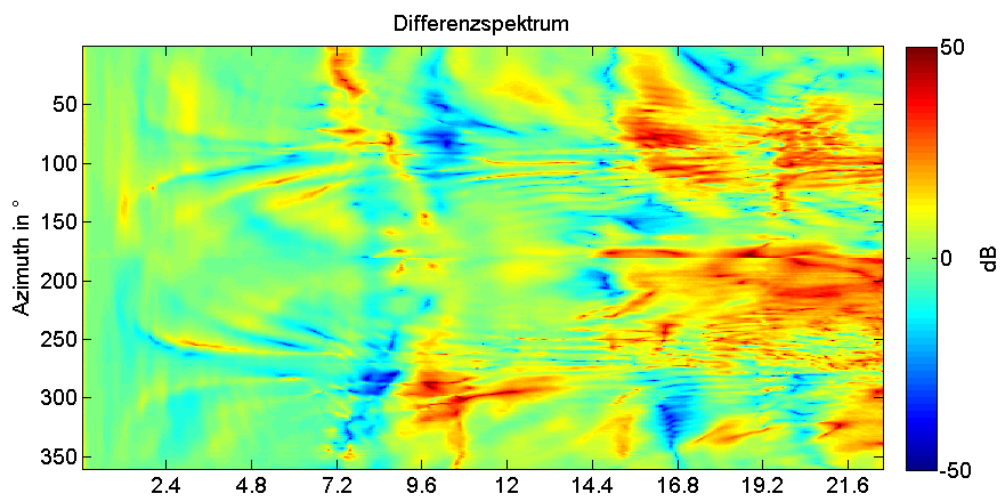


Abbildung 31: Differenzspektrum zwischen Proband (ltu.mat) und Kunstkopf

Beim genaueren Betrachten von den beiden einzelnen Spektren ist noch ein weiterer eindeutiger Unterschied zwischen Kunstkopf und Probandendaten sichtbar. Dieser wurde in Abbildung 32 visualisiert.

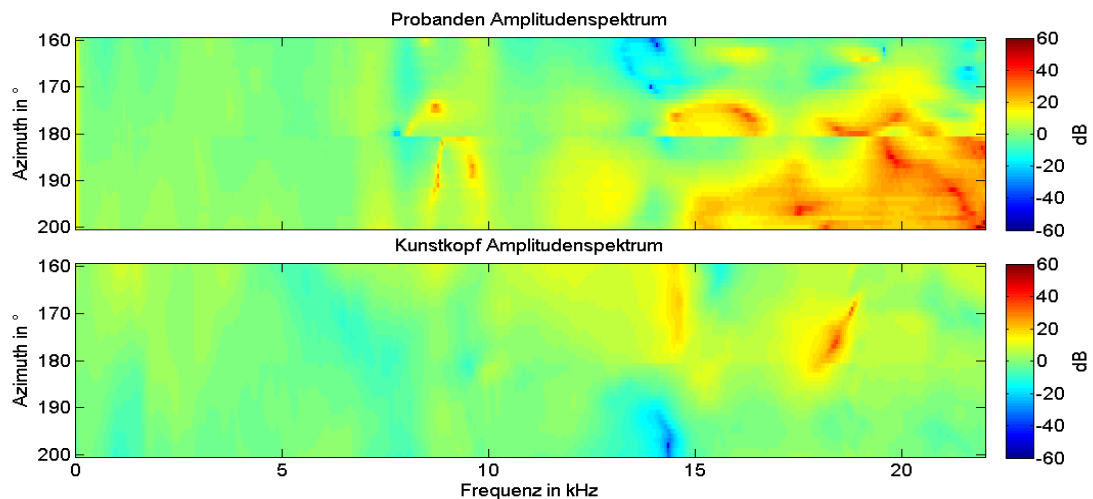


Abbildung 32: Amplitudendifferenzvergleich von Proband (ltu.mat) und Kunstkopf von 160° bis 200°

In diesem vergrößerten Abschnitt sieht man einen Versatz bei 180°. Dieser Absatz ist bei allen Probanden sehr deutlich bei 90°, 180° und 270° zu erkennen. Er verdeutlicht das Ende beziehungsweise den Anfang einer Messung, da mit vier Lautsprechern jeweils 90° aufgenommen wurden. Ein kaum merklicher Unterschied ist auch beim Kunstkopf vorhanden, was darauf zurückzuführen ist, dass die Lautsprecher nicht hundertprozentig korrekt ausgerichtet wurden. Der deutlich sichtbare Versatz beim Probanden ist auf den Fakt zurückzuführen, dass der Kopf der vermessenen Person sich nicht im Drehmittelpunkt befand und der Proband seinen Kopf am Ende der Messung nicht in der selben Position hatte wie zu Beginn.

Restliche Fehler, die sich vor allem in der sprunghaften Veränderung der ITD zeigen, sind auf allgemeine Durchführungsfehler zurückzuführen, die unbewusst vom Probanden verursacht wurden. Solche Fehler sind Räuspern, Husten, Schlucken, Änderung der Sitzposition oder Halterung, sowie die Schultern oder den Kopf im Laufe der Messung zu senken. Auch Herzschlaggeräusche im Ohr aufgrund von Aufregung des Probanden können eine Fehlerquelle sein. Ein weiterer möglicher Fehler ist, dass der Proband sich

unbewusst auf einen Lautsprecher fixiert und dadurch den Kopf dreht. Möglich ist auch, dass eine der Sonden im Ohr des Probanden verrutscht.

5. Fazit und Ausblick

Bei der Aufnahme von kopfbezogenen Übertragungsfunktionen müssen Vorüberlegungen bezüglich der Herangehensweise getroffen werden. Diese Arbeit beschäftigt sich mit der Herausforderung, eine solche Aufnahme durchzuführen, und zwar so, dass sie auch für individualisierte Messungen genutzt werden kann. Im Rahmen dieser Arbeit wurden die Übertragungsfunktionen beispielhaft ausgewählter Probandenmessungen mit denen eines Kunstkopfes verglichen. Die Aufnahmebedingungen sind die gleichen. Die HRTF - Aufnahme des Kunstkopfes lieferte die Referenz für den Vergleich mit den Probanden. Um die Aufnahme zugunsten der Probanden zu verkürzen und den Stressfaktor gering zu halten, wurde mit vier Lautsprechern gearbeitet. Somit mussten die Probanden nur um 90° gedreht werden. Davon wird in zukünftigen HRTF - Messungen abgeraten. Die korrekte Ausrichtung und Entzerrung aller Lautsprecher bietet großes Fehlerpotential. Der Fehler ist auch in der Auswertung deutlich zu erkennen.

Die grafische Auswertung des Vergleichs zeigt, dass die Pegeldifferenzen eine Figur erzeugen, welche bei Probanden und Kunstkopf auffallend ähnlich ist. Inwiefern die verschiedenen HRTF - Aufnahmen bei unterschiedlichen Hörern verwendet werden können, war nicht Gegenstand dieser Arbeit. Zur Untersuchung dieser Frage bietet sich ein Hörtest an.

Die HRTF - Aufnahme des Kunstkopfes wurde auch als Ausgangspunkt für die Auswahl der Methode zur Errechnung der Laufzeitdifferenz genutzt. Es wurden zwei Methoden der Laufzeitdifferenz verglichen. Die Methode mit den geringsten Abweichungen wurde für diese Arbeit ausgewählt. Das Ergebnis zeigt jedoch, dass der Erfolg der Anwendung dieser Methode bei den individualisierten HRTF's geringer ausfällt, als bei den Kunstkopf - HRTF's. Es wird aufgrund der fehlerbehafteten Ergebnisse der Laufzeitberechnung vorgeschlagen, eine Kombination aus den beiden verwendeten Methoden zu testen, da bei beiden als separate Anwendungen mit Abweichungen zu rechnen ist. Dies kann ein Thema für eine weitere Untersuchung sein.

Die Arbeit soll Aufschluss über die Verwendung einer Minimalphasenapproximation für die HRTF - Daten geben. Die Datenbank des

Kunstkopfes beinhaltet aus dem Grund auch die Minimalphasenversion der Aufnahme. Das minimalphasige Amplitudenspektrum kann nun mit individuellen Laufzeitdifferenzen verknüpft werden, um eine globale ITD - Anpassung zu ermöglichen. Auch die Phase ändert sich mit dieser Anpassung individuell. Die grafische und analytische Fehleranalyse dieser Daten im Vergleich zum Original zeigt, dass die Minimalphasenberechnung gut funktioniert und nur geringe Abweichungen im Amplitudenspektrum zum Original bestehen. Kistler und Wighman bestätigen, dass Lokalisationsversuche mit diesen Modellen sehr ähnliche Ergebnisse wie Freifeldversuche liefern [Kistler, 1992]. Das erlaubt die weitere Verwendung der Minimalphasenanwendung in solchen Zusammenhängen.

Die Überlegung der Länge der Impulsantworten und damit verbunden die Erstellung einer geeigneten Fensterfunktion war Teil dieser Arbeit. Diese Überlegungen waren sehr wichtig, um eine Vorstellung von der korrekten Länge und dem Erscheinungsbild einer Impulsantwort für dieses Thema zu erlangen.

6. Abbildungsverzeichnis

Abbildung 1: Kopfbezogenes, dreidimensionales Koordinatensystem mit Azimuthwinkel φ und Elevationswinkel ϑ [Fels, 2008].....	3
Abbildung 2: Prinzipskizze zur Bestimmung der Laufzeitdifferenz [Kremer, 2012].....	4
Abbildung 3: Relevanz der Differenzen anhand der Duplex Theorie [Breebart, 2007].....	5
Abbildung 4: Demonstration der Breite eines akustischen Objektes anhand einer Stereolautsprecheranordnung, Breite erhöht sich (1-3), während sich die IC zwischen den beiden Lautsprechern verringert. [Breebart, 2007].....	7
Abbildung 5: Richtungslokalisation in der Medianebene [Hübner, 2011].....	8
Abbildung 6: Grundform einer HRTF. [Fedke, 2007, S.5].....	10
Abbildung 7: Paarbeispiele zweier HRIR (links) und HRTF (rechts), aufgenommen in einem schallarmen Raum. [Datenbank Oldenburg].....	11
Abbildung 8: Binaurale Aufnahme mittels Mikrofonen im Kunstkopfgehörgang (links) welche über Kopfhörer wiedergegeben werden (rechts). [Breebart, 2007, S.10].....	14
Abbildung 9: Mikrofonanordnung für Laufzeitstereophonie.[Bartlett,1991, S.19].....	16
Abbildung 10: Mikrofonanordnung für Pegeldifferenzstereophonie.[Bartlett, 1991, S.21].....	16
Abbildung 11: Grundsätzlicher Aufbau der Äquivalenzstereophonie, z.B ORTF. [Bartlett,1991,S.23].....	16
Abbildung 12: Aufbau einer Mikrofonierung mittels Jecklin-Scheibe mit seitlichem Schalleinfall. [Curd, 2011]	17
Abbildung 13: Beispiel eines Kugelflächenmikrofons.[Ederhof, 2004, S.212]... ..	18
Abbildung 14: Beispiel eines Kunstkopfes von der Firma Neumann (KU100). [Fedke].....	18
Abbildung 15: G.R.A.S Kemar Head and Torso Simulator BA45. [verwendeter Kunstkopf].....	19
Abbildung 16: Beispielhafter Sinussweep im Zeitbereich. [eigene Simulation]..	21
Abbildung 17: Ausschnitt eines MLS - Rauschens der Ordnung 8. [eigene Simulation].....	22
Abbildung 18: Grundsätzlicher Messaufbau im Schallarmen Raum, links: Prinzipskizze, rechts: Fotodarstellung.....	25
Abbildung 19: Laser zur korrekten Messpositionierung.....	25
Abbildung 20: Beispielimpulsantwort der gemessenen Datenbank und passende Fensterung.....	26
Abbildung 21: Vergleich des Ergebnisses beider Methoden zur ITD-Berechnung (links: Maximalwertsuche, rechts: Anstiegssuche (15 %)). [eigene Messung].	29
Abbildung 22: Oben: Impulsantwort eines Systems. Unten: Minimalphasenreduzierte Impulsantwort.[Datenbank Oldenburg].....	30
Abbildung 23: Abfolge der Signalverarbeitung.....	31
Abbildung 24: Betrachtung des Fehlers im Frequenzbereich der Minimalphasendaten in Bezug zum Original.	32
Abbildung 25: Beispiel eines rekonstruierten und eines Originalen HRTF - Paares (10° Azimut).....	32
Abbildung 26: Vergleich der Phase und des Amplitudenspektrums von rekonstruierten und originalen Daten. (linkes Ohr, 10° Azimut).....	33

Abbildung 27: Amplitudenspektrum der Differenz beider Ohren als Farbintensitätsdarstellung.....	33
Abbildung 28: Laufzeitvergleich zwischen Proband (beer.mat) und Kunstkopf.	34
Abbildung 29: Vergleich der Zeitfunktion von Proband (beer.mat) und Kunstkopf für einen Azimut von 254°	35
Abbildung 30: Amplitudendifferenzvergleich von Proband (ltu.mat) und Kunstkopf.....	36
Abbildung 31: Differenzspektrum zwischen Proband (ltu.mat) und Kunstkopf.	36
Abbildung 32: Amplitudendifferenzvergleich von Proband (ltu.mat) und Kunstkopf von 160° bis 200°	37

7. Literaturverzeichnis

- [Bartlett, 1991] B. Bartlett, *Stereo microphone techniques*. Boston: Focal Press, 1991
- [Blauert, 1997] J. Blauert, *Spatial hearing: The psychophysics of human sound localization*. Cambridge, Mass: MIT Press, 1997.
- [Blommer, 1997] M. Blommer and G. Wakefield, "Pole-zero approximations for head-related transfer functions using a logarithmic error criterion," *IEEE Transactions on Speech and Audio Processing*, vol. 5, no. 3, pp. 278–287, <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=568734>, 1997.
- [Breebart, 2007] J. Breebaart and C. Faller, *Spatial audio processing: MPEG surround and other applications*. Chichester, West Sussex, England ;, Hoboken, NJ: John Wiley & Sons, 2007.
- [Carlile, 2000] S. Carlile, C. Jin, and V. van Raad, "Continuous virtual auditory space using HRTF interpolation: acoustic and psychophysical errors," in *IEEE PacificRim Conference on Multimedia*, 2000, pp. 220-223.
- [Cipic, 2001] V. Algazi, R. Duda, D. Thompson, and C. Avendano, "The CIPIC HRTF database," in *Applications of Signal Processing to Audio and Acoustics, 2001 IEEE Workshop on the*, 2001, pp. 99–102.
- [Curd, 2011] O. Curd. (2011, October 17). *Trennkörper Mikrofonsysteme, Audiovisuelle Medien: Ton Seminar SS/2010f*. [Online] Available: <http://www.hdm-stuttgart.de/~curdt/Trennk%C3%B6rper-Mikrofonsysteme.pdf>
- [Ederhof, 2004] A. Ederhof, *Das Mikrofonsbuch: Optimaler Einsatz im Studio und auf der Bühne*. München: GC Carstensen, 2004.
- [Faller, 2001] C. Faller and F. Baumgarte, "Efficient representation of spatial audio using perceptual parametrization," in *Applications of Signal Processing to Audio and Acoustics, 2001 IEEE Workshop on the*, 2001, pp. 199–202.
- [Faller, 2004] C. Faller and J. Merimaa, "Source localization in complex listening situations: Selection of binaural cues based on interaural coherence," *J. Acoust. Soc. Am*, vol. 116, no. 5, p. 3075, 2004.

- [Farina, 2000]** A. Farina, Ed, *Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique*, 2000.
- [Fedtke, 2007]** T. Fedtke et. al.: „Kunstkopftechnik – Eine Bestandsaufnahme“ Mitteilung aus dem Normenausschuss „Psychoakustische Messtechnik“ in *Acta Acustica/Acustica* 93:1.
- [Fels, 2008]** J. Fels, *From children to adults: How binaural cues and ear canal impedances grow*. Dissertation, Rheinisch-Westf. Technische Hochschule Aachen 2008.
- [Four, 2008]** Four Audio GmbH, *ELF User Manual*. Halle, Germany.
- [Hübner, 2011]** Hüber and J.-F. Timm. (2011, October 10). *Darstellung der Richtungswahrnehmung bestimmter Frequenzbereiche*. [Online] Available: <http://www.computerbase.de/bildstrecke/23087/20/>.
- [Kemar, 2008]** G.R.A.S Sound & Vibration A/S, *Data Sheet for 45BA*. Holte, DK.
- [Kistler, 1992]** Kistler, D.J. and Wightman, F.L., “A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction,” *J. Acoust. Soc. Am.*, Vol. 91, Wight”, EL.. and Kistler, 1992, p ~1.63 7-1647
- [Kremer, 2012]** M. Kremer. (2011, October 12). *Ars Auditus*. [Online] Available: http://www.dasp.uni-wuppertal.de/ars_auditus/ueberfunktion/funk24.htm.
- [Kulkarni, 2002]** A. Kulkarni, S. Isabelle, and H. Colburn, “On the minimum-phase approximation of head-related transfer functions,” in *Applications of Signal Processing to Audio and Acoustics, 1995, IEEE ASSP Workshop on*, 1995, pp. 84–87.
- [Lindau, 2006]** A. Lindau: „Ein Instrument zur softwaregestützten Messung binauraler Raumimpulsantworten in mehreren Freiheitsgraden“, Magisterarbeit, Creative Commons, 2006.
- [Middlebrooks, 2002]** E. A. Macpherson and J. C. Middlebrooks, “Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited,” *J. Acoust. Soc. Am*, vol. 111, no. 5, p. 2219, 2002.

- [Møller, 1996]** H. Møller, M. F. Sørensen, C. B. Jensen, and D. Hammershøi, "Binaural Technique: Do We Need Individual Recordings?," *J. Audio Eng. Soc.*, vol. 44, no. 6, pp. 451-469, <http://www.aes.org/e-lib/browse.cfm?elib=7897>, 1996.
- [Möser, 2009]** M. Möser, Ed, *Messtechnik der Akustik*. Berlin: Springer, 2010.
- [Oldenburg, 2009]** H. Kayser, S. D. Ewert, J. Anemüller, T. Rohdenburg, V. Hohmann, and B. Kollmeier, "Database of multichannel in-ear and behind-the-ear head-related and binaural room impulse responses," (eng), *EURASIP journal on advances in signal processing*, vol. 2009.
- [Park, 2005]** S. Hwang, Y. Park, and Y. Park, "Sound Source Localization using HRTF database," *ICCAS*, p. 5, <http://vasl.kaist.ac.kr/sdakupload/paper/sungmokHwang.pdf>, 2005.
- [Vanderkooy, 1989]** D. D. Rife and J. Vanderkooy, "Transfer Function Measurements using Maximum Length Sequences," *J. Audio Eng. soc.*, no. 6, pp. 419-444, 1989.
- [Yost, 1994]** W. A. Yost, *Fundamentals of hearing: An introduction*, 3rd ed. San Diego, Calif: Acad. Press, 1994.
- [Zotkin, 2003]** D. Zotkin, J. Hwang, R. Duraiswaini, and L. Davis, "HRTF personalization using anthropometric measurements," in *Applications of Signal Processing to Audio and Acoustics*, 2003 IEEE Workshop on, 2003, pp. 157-160.

8. Abkürzungs - und Symbolverzeichnis

Symbol oder Abkürzung	Bedeutung
°C	Grad Celsius
c	Schallgeschwindigkeit in Luft bei 20 °C (343 m/s)
cm	Zentimeter
dB	Dezibel
Fs	Abtastrate (48000 Hz)
Hz	Hertz
HRIR	Head-related impulse response
HRTF	Head-related transfer function
ILD	Interaural Level Difference
IPD	Interaural Phase Difference
ITD	interaural time difference
kHz	Kilohertz
m/s	Meter pro Sekunde
ms	Millisekunde
µs	Mikrosekunde
s	Weg bis zum Kopfmikrofon in Meter
t	Zeit
vz	Verzögerung in Samples

Ehrenwörtliche Erklärung

Hiermit versichere ich, die vorliegende Arbeit selbstständig und ohne fremde Hilfe angefertigt zu haben. Die verwendete Literatur und sonstige Hilfsmittel sind vollständig angegeben.

Ilmenau, 09. März 2012

M. Hellmich

Thesen

- Eine Minimalphasenapproximation stellt für die Wahrnehmung eine gute Näherung an die Originaldaten dar.
- Messungen mit mehreren Lautsprechern sind an den Grenzübergängen stark fehlerbehaftet.
- Die Anpassung der einzelnen Lautsprecher ist sehr aufwendig.
- Ein Kunstkopf stellt eine praktikable Referenz als Grundlage für Probandenmessungen dar.
- Keine der untersuchten Methoden zur Laufzeitdifferenzberechnung sind für Probandenmessungen optimal geeignet.